

Issue: 33

SCOPUS

UNIVERSIDAD POLITÉCNICA SALESIANA ECUADOR

pISSN: 1390-650X

eISSN: 1390-860X

January / June 2025

INGENIUS

Revista de Ciencia y Tecnología



Q4
SJR Journal
Rank

SUBJECT AREA AND CATEGORY
Engineering
(miscellaneous)

- Determination of optimal formats for digital image

Pag. 9

- Application of manets as a communication system for sustainable mobility

Pag. 27

- A comprehensive evaluation of AI techniques for air quality index prediction: RNNS and transformers

Pag. 60

- Comparative analysis of flow patterns in off-design planar and conical nozzle

Pag. 115

Indexada en: **SCOPUS**

INGENIUS

INGENIUS • Issue 33 • January/June 2025. Journal of Science and Technology of the Universidad Politécnica Salesiana of Ecuador. Publication dedicated to studies related to the Sciences of Mechanical Engineering, Electrical Engineering, Electronic Engineering, Mechatronic Engineering, Systems Engineering and Industrial Engineering.

Editors Board

RAFAEL ANTONIO BALART GIMENO, PHD, Universidad Politécnica de Valencia, España – Editor-in-chief.

JOHN IGNACIO CALLE SIGUENCIA, PHD, Universidad Politécnica Salesiana, Ecuador – Editor-in-chief.

MARIELA CERRADA LOZADA, PHD, Universidad Politécnica Salesiana, Ecuador – Associate Editor.

TEODIANO FREIRE BASTOS FILHO, PHD, (Universidade Federal do Espírito Santo, Brasil – Associate Editor.

MARLON XAVIER QUINDE ABRIL, MSC, Universidad Politécnica Salesiana, Ecuador – Associate Editor.

Scientific board

JUAN LÓPEZ MARTÍNEZ, PHD, Universidad Politécnica de Valencia, España.

ELENA FORTUNATI, PHD, Universidad de Perugia, Italia.

GUSTAVO ROVELO RUIZ, PHD, Hasselt University, Diepenbeek, Bélgica.

FRANKLIN GAVILANEZ ALVAREZ, PHD, American University, Estados Unidos.

PIEDAD GAÑAN ROJO, PHD, Universidad Pontificia Bolivariana, Colombia.

JOSÉ ALEX RESTREPO, PHD, Universidad Simón Bolívar, Venezuela.

SERGIO LUJAN MORA, PHD, Universidad de Alicante, España.

MARTHA ZEQUERA DÍAZ, PHD, Pontificia Universidad Javeriana, Colombia.

GROVER ZURITA, PHD, Universidad Privada Boliviana, Bolivia.

VLADIMIR ROBLES, PHD, Universidad Politécnica Salesiana, Ecuador.

GERMÁN ARÉVALO, PHD, Universidad Politécnica Salesiana, Ecuador.

WILBERT AGUILAR, PHD, Universidad de las Fuerzas Armadas, ESPE, Ecuador.

JACK BRAVO TORRES, PHD, Universidad Politécnica Salesiana, Ecuador.

WALTER OROZCO, PHD, Universidad Politécnica Salesiana, Ecuador.

MARIELA CERRADA, PHD, Universidad Politécnica Salesiana, Ecuador.

JULIO CÉSAR VIOLA, PHD, Universidad Politécnica Salesiana, Ecuador.

SERGIO GAMBOA SÁNCHEZ, PHD, Universidad Nacional Autónoma de México, México.

ROGER ABDÓN BUSTAMANTE PLAZA, PHD, Universidad de Chile, Chile.

CHRISTIAN BLUM, PHD, Consejo Superior de Investigaciones Científicas, España.

SILVIA NOEMI SCHIAFFINO, PHD, Universidad Nacional del Centro de la Provincia de Buenos Aires, Argentina.

ANALÍA ADRIANA AMANDI, PHD, Universidad Nacional del Centro de la Provincia de Buenos Aires, Argentina.

RUBÉN DE JESÚS MEDINA MOLINA, PHD,

Universidad de Los Andes, Venezuela.

JOHNNY JOSUÉ BULLÓN TORREALBA, PHD, Universidad de Los Andes, Venezuela.

RODRIGO PALMA HILLERNS, PHD, Universidad de Chile, Chile.

GERARDO ESPINOZA PÉREZ, PHD, Universidad Nacional Autónoma de México, México.

ALEXANDRE MENDES ABRÃO, PHD, Universidad Federal de Minas Gerais, Brasil.

KAMLA ABDEL RADI ISMAIL, PHD, Universidad Estatal de Campinas Unicamp, Brasil.

ARNALDO DA SILVA, PHD, Universidad Estatal de Campinas Unicamp, Brasil.

ÁLVARO ROCHA, PHD, Universidad de Coimbra, Portugal.

JOSÉ ANTENOR POMILIO, PHD, Universidad Estatal de Campinas Unicamp, Brasil.

LUIS PAULO REIS, PHD, Universidad de Minho, Portugal.

LUÍS FERNANDES, PHD, Escuela Superior Náutica Infante d. Henrique, Portugal.

ANÍBAL TRAÇA DE ALMEIDA, PHD, Universidad de Coimbra, Portugal.

JORGE SÁ SILVA, PHD, Universidad de Coimbra, Portugal.

PEDRO MANUEL SOARES MOURA, PHD, Universidad de Coimbra, Portugal.

SÉRGIO MANUEL RODRIGUES LOPES, PHD, Universidad de Coimbra, Portugal.

RICARDO MADEIRA SOARES BRANCO, PHD, Universidad de Coimbra, Portugal.

CARLOS ALEXANDRE BENTO CAPELA, PHD, Universidad de Coimbra, Portugal.

FILIFE ARAUJO, PHD, Universidad de Coimbra, Portugal.

LUIS MANUEL GUERRA SILVA ROSA, PHD, Universidad de Lisboa, Portugal.

HÉLDER DE JESUS FERNANDES, PUGA, PHD, Universidad de Minho, Portugal.

FILIFE SAMUEL, PEREIRA DA SILVA, PHD, Universidad de Minho, Portugal.

CÉSAR SEQUEIRA, PHD, Universidad de Lisboa, Portugal.

JOSÉ TEIXEIRA ESTÊVÃO FERREIRA, PHD,

Universidad de Coimbra, Portugal.

NUNO LARANJEIRO, PHD, Universidad de Coimbra, Portugal.

LUÍS AMARAL, PHD, Universidad de Lisboa, Portugal.

JORGE HENRIQUES, PHD, Universidad de Coimbra, Portugal.

WILLIAM IPANAQUE, PHD, Universidad de Piura, Perú.

LORENZO LEIJA SALAS, PHD, Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional, México.

VALERI KONTOROVICH MAZOVER, PHD, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, México.

ALEJANDRO ÁVILA GARCÍA, PHD, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, México.

PAOLO BELLAVISTA, PHD, Universidad de Bologna, Italia.

CARLOS RUBIO, PhD, Centro de Ingeniería y Desarrollo Industrial, México.

FERNANDO HERNÁNDEZ SÁNCHEZ, PhD, Centro de Investigación Científica de Yucatán, México.

EMILIO MUÑOZ SANDOVAL, PhD, Instituto Potosino de Investigación Científica y Tecnológica, México.

YASUHIRO MATSUMOTO KUWABARA, PhD, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, México.

DAVID ZUMOFFEN, PhD, Centro Internacional Franco Argentino de Ciencias de la Información y de Sistemas, Argentina.

VICENTE RODRÍGUEZ GONZÁLEZ, PhD, Instituto Potosino de Investigación Científica y Tecnológica, México.

ALEJANDRO RODRÍGUEZ ÁNGELES, PhD, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, México.

ALISTAIR BORTHWICK, PhD, Universidad de Edimburgo, Reino Unido.

Reviewers board

FEDERICO DOMINGUEZ, PHD, Escuela Superior Politécnica del Litoral, Ecuador.

ENRIQUE CARRERA, PHD, Universidad de las Fuerzas Armadas, ESPE, Ecuador.

ANDRÉS TELLO, MSc, Universidad de Cuenca, Ecuador.

CRISTIAN GARCÍA BAUZA, PHD, Universidad Nacional del Centro de la Provincia de Buenos Aires, Argentina.

OSVALDO AÑÓ, PHD, Universidad Nacional de San Juan, Argentina.

THALÍA SAN ANTONIO, PHD, Universidad Técnica de Ambato, Ecuador.

VICTOR SAQUICELA, PHD, Universidad de Cuenca, Ecuador.

GONZALO OLMEDO, PHD, Universidad de las Fuerzas Armadas, ESPE, Ecuador.

ROMÁN LARA, PHD, Universidad de las Fuerzas Armadas, ESPE, Ecuador.

GUILLERMO SORIANO, PHD, Escuela Superior Politécnica del Litoral, Ecuador.

MARÍA FERNANDA GRANDA, PHD, Universidad de Cuenca, Ecuador.

RICARDO CAYSSIALS, PHD, Universidad Tecnológica Nacional, Argentina.

LEONARDO SOLAQUE GUZMAN, PHD, Universidad Militar Nueva Granada, Colombia.

JOSÉ DI PAOLO, PHD, Universidad Nacional de Entre Ríos, Argentina.

ASTRID RUBIANO FONSECA, PHD, Universidad Militar Nueva Granada, Colombia.

ROBINSON JIMÉNEZ, PHD, Universidad Militar Nueva Granada, Colombia.

ALFONSO ZOZAYA, PHD, Universidad de Carabobo, Venezuela.

MAURICIO MAULEDOUX, PHD, Universidad Militar Nueva Granada, Colombia.

LUIS MEDINA, PHD, Universidad Simón Bolívar, Venezuela.

ERNESTO CUADROS-VARGAS, PHD, Universidad Católica San Pablo, Perú.

SAMUEL SEPÚLVEDA CUEVAS, PHD, Universidad de la Frontera, Chile.

CARLOS CARES, PHD, Universidad de la Frontera, Chile.

RAFAEL SOTELO, PHD, Universidad de Montevideo, Uruguay.

OMAR LOPEZ, PHD, Universidad de Los Andes, Colombia.

JOB FLORES-GODOY, PHD, Universidad Católica del Uruguay, Uruguay.

LUIS MARIO MATEUS, PHD, Universidad de los Andes, Colombia.

AMADEO ARGÜELLES CRUZ, PHD, Instituto Politécnico Nacional, México.

SANTIAGO BENTANCOURT PARRA, PHD, Universidad Pontificia Bolivariana, Colombia.

GERMÁN ZAPATA, PHD, Universidad Nacio-

nal de Colombia, Colombia.

PEDRO GARCÍA, PHD, Universidad Autónoma de Barcelona, España.

ARTURO CONDE ENRÍQUEZ, PHD, Universidad Autónoma de Nuevo León, México.

ALBERTO CAVAZOS GONZÁLEZ, PHD, Universidad Autónoma de Nuevo León, México.

ERNESTO VÁZQUEZ MARTÍNEZ, PHD, Universidad Autónoma de Nuevo León, México.

MIGUEL DÍAZ RODRIGUEZ, PHD, Universidad de Los Andes, Venezuela.

EFRAÍN ALCORTA GARCÍA, PHD, Universidad Autónoma de Nuevo León, México.

LUIS CHIRINOS GARCIA, PHD, Pontificia Universidad Católica de Perú, Perú.

OSCAR AVILÉS, PHD, Universidad Militar Nueva Granada, Colombia.

DORA MARTÍNEZ DELGADO, PHD, Universidad Autónoma de Nuevo León, México.

DAVID OJEDA, PHD, Universidad Técnica del Norte, Ecuador.

IRENE BEATRÍZ STEINMANN, PHD, Universidad Tecnológica Nacional, Argentina.

MARIO SERRANO, Universidad Nacional de San Juan, Argentina.

CORNELIO POSADAS CASTILLO, PHD, Universidad Autónoma Nuevo León, México.

MARIO ALBERTO RIOS MESIAS, PHD, Universidad de Los Andes, Colombia.

YUDITH CARDINALE VILLARREAL, PHD, Universidad Simón Bolívar, Venezuela.

JOSE EDUARDO OCHOA LUNA, PHD, Universidad Católica San Pablo, Perú.

DANTE ANGEL ELIAS GIORDANO, PHD, Pontificia Universidad Católica de Perú, Perú.

MANUEL PELAEZ SAMANIEGO, PHD, Universidad de Cuenca, Ecuador.

JUAN ESPINOZA ABAD, PHD, Universidad de Cuenca, Ecuador.

PIETRO CODARA, PHD, Universidad de Milan, Italia.

ALBERTO SORIA, PHD, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, México.

JOSÉ M. ALLER, PHD, Universidad Politécnica Salesiana, Ecuador.

FERNEY AMAYA F., PHD, Universidad Pontificia Bolivariana, Medellín, Colombia.

SANTIAGO ARANGO ARAMBURO, PHD, Universidad Nacional de Colombia, Colombia.

DIEGO ARCOS-AVILÉS, PHD, Universidad de las Fuerzas Armadas, ESPE, Ecuador.

PABLO AREVALO, PHD, Universidad Politécnica Salesiana, Ecuador.

ROBERTO BELTRAN, MSc, Universidad de las Fuerzas Armadas, ESPE, Ecuador.

LEONARDO BETANCUR, PHD, Universidad Pontificia Bolivariana, Medellín, Colombia.

ROBERTO GAMBOA, PHD, Universidad de Lisboa, Portugal.

PAULO LOPES DOS SANTOS, PHD, Universidad do Porto, Portugal.

PEDRO ANDRÉ DIAS PRATES, PHD, Universidad de Coimbra, Portugal.

JOSÉ MANUEL TORRES FARINHA, PHD, Universidad de Coimbra, Portugal.

CELSO DE ALMEIDA, PHD, Universidad Estatal de Campinas Unicamp, Brasil.

RAMON MOLINA VALLE, PHD, Universidad Federal de Minas Gerais, Brasil.

CRISTINA NADER VASCONCELOS, PHD, Universidad Federal Fluminense, Brasil.

JOÃO M. FERREIRA CALADO, PHD, Universidad de Lisboa, Portugal.

GUILHERME LUZ TORTORELLA, PHD, Universidad Federal de Santa Catarina, Brasil.

MAURO E. BENEDET, PHD, Universidad Federal de Santa Catarina, Brasil.

ARTEMIS MARTI CESCHIN, PHD, Universidade de Brasilia, Brasil.

GILMAR BARRETO, PHD, Universidad Estatal de Campinas Unicamp, Brasil.

RICARDO EMILIO F. QUEVEDO NOGUEIRA, PHD, Universidad Federal de Ceará, Brasil.

WESLEY LUIZ DA SILVA ASSIS, PHD, Universidad Federal Fluminense, Brasil.

ANA P. MARTINAZZO, PHD, Universidad Federal Fluminense, Brasil.

JORGE BERNARDINO, PHD, Universidad de Coimbra, Portugal.

LUIS GERALDO PEDROSO MELONI, PHD, Universidad Estatal de Campinas Unicamp, Brasil.

FACUNDO ALMERAYA CALDERÓN, PHD, Universidad Autónoma de Nuevo León, México.

FREDDY VILLAO QUEZADA, PHD, Escuela Superior Politécnica del Litoral, Ecuador.

JOSE MANRIQUE SILUPU, MSc, Universidad de Piura, Perú.

GERMÁN ARIEL SALAZAR, PHD, Instituto de Investigaciones en Energía no Convencional, Argentina.

JOSÉ MAHOMAR JANANÍAS, PHD, Universidad del BIOBIO, Chile.

ARNALDO JÉLVEZ CAAMAÑO, PHD, Universidad del BIOBIO, Chile.

JORGE ANDRÉS URIBE, MSc, Centro de Ingeniería y Desarrollo Industrial, México.

RICARDO BELTRAN, PHD, Centro de Investigación en Materiales Avanzados, México.

ADI CORRALES, MSc, Centro de Ingeniería y Desarrollo Industrial, México.

JORGE URIBE CALDERÓN, PHD, Centro de Investigación Científica de Yucatán, México.

JOSÉ TRINIDAD HOLGUÍN MOMACA, MSc, Centro de Investigación en Materiales Avan-

zados, México.

JUAN MANUEL ALVARADO OROZCO, PhD, Centro de Ingeniería y Desarrollo Industrial, México.

ARNALDO JÉLVEZ CAAMAÑO, PHD, Universidad del BIOBIO, Chile.

JAVIER MURILLO, PHD, Centro Internacional Franco Argentino de Ciencias de la Información y de Sistemas, Argentina.

LUCAS DANIEL TERRISSI, PHD, Universidad Nacional de Rosario, Argentina.

RENE VINICIO SANCHEZ LOJA, MSC, Universidad Politécnica Salesiana, Ecuador.

FREDDY LEONARDO BUENO PALOMEQUE, MSC, Universidad Politécnica Salesiana, Ecuador.

DIEGO CABRERA MENDIETA, MSC, Universidad Politécnica Salesiana, Ecuador.

EDWUIN JESUS CARRASQUERO, PHD, Universidad Técnica de Machala, Ecuador.

CARLOS MAURICIO CARRILLO ROSERO, MSC, Universidad Técnica de Ambato, Ecuador.

DIEGO CARRION GALARZA, MSC, Universidad Politécnica Salesiana, Ecuador.

CARMEN CELI SANCHEZ, MSC, Universidad Politécnica Salesiana, Ecuador.

DIEGO CHACON TROYA, MSC, Universidad Politécnica Salesiana, Ecuador.

PAUL CHASI, MSC, Universidad Politécnica Salesiana, Ecuador.

JUAN CHICA, MSC, Universidad Politécnica Salesiana, Ecuador.

DIEGO MARCELO CORDERO GUZMÁN, MSC, Universidad Católica de Cuenca, Ecuador.

LUIS JAVIER CRUZ, PHD, Universidad Pontificia Bolivariana, Medellín, Colombia.

FABRICIO ESTEBAN ESPINOZA MOLINA, MSC, Universidad Politécnica Salesiana, Ecuador.

JORGE FAJARDO SEMINARIO, MSC, Universidad Politécnica Salesiana, Ecuador.

PATRICIA FERNANDEZ MORALES, PHD, Universidad Pontificia Bolivariana, Medellín, Colombia.

MARCELO FLORES VAZQUEZ, MSC, Universidad Politécnica Salesiana, Ecuador.

CARLOS FLORES VÁZQUEZ, MSC, Universidad Católica de Cuenca, Ecuador.

CARLOS FRANCO CARDONA, PHD, Universidad Nacional de Colombia, Colombia.

CRISTIAN GARCÍA GARCÍA, MSC, Universidad Politécnica Salesiana, Ecuador.

TEONILA GARCÍA ZAPATA, PHD, Universidad Nacional Mayor de San Marcos, Perú.

LUIS GARZÓN MÑOZ, PHD, Universidad Politécnica Salesiana, Ecuador.

NATALIA GONZALEZ ALVAREZ, MSC, Universidad Politécnica Salesiana, Ecuador.

ERNESTO GRANADO, PHD, Universidad Simón Bolívar, Venezuela.

ADRIANA DEL PILAR GUAMAN, MSC, Universidad Politécnica Salesiana, Ecuador.

JUAN INGA ORTEGA, MSC, Universidad Politécnica Salesiana, Ecuador.

ESTEBAN INGA ORTEGA, PHD, Universidad Politécnica Salesiana, Ecuador.

PAOLA INGAVÉLEZ, MSC, Universidad Politécnica Salesiana, Ecuador.

CESAR ISAZA ROLDAN, PHD, Universidad Pontificia Bolivariana.

NELSON JARA COBOS, MSC, Universidad Politécnica Salesiana, Ecuador.

RUBEN JERVES, MSC, Universidad Politécnica Salesiana, Ecuador.

VICTOR RAMON LEAL, PHD, Investigador de PDVSA, Venezuela

GABRIEL LEON, MSC, Universidad Politécnica Salesiana, Ecuador.

EDILBERTO LLANES, PHD, Universidad Internacional SEK, Ecuador.

LUIS LÓPEZ, MSC, Universidad Politécnica Salesiana, Ecuador.

CARLOS MAFLA YÉPEZ, MSC, Universidad Técnica del Norte, Ecuador.

HADER MARTÍNEZ, PHD, Universidad Pontificia Bolivariana, Medellín, Colombia

JAVIER MARTÍNEZ, PHD, Instituto Nacional de Eficiencia Energética y Energías Renovables, Ecuador.

ALEX MAYORGA, MSC, Universidad Técnica de Ambato, Ecuador.

JIMMY MOLINA, MSC, Universidad Técnica de Machala, Ecuador.

ANDRES MONTERO, PHD, Universidad de Cuenca, Ecuador.

VICENTE MORALES, MSC, Universidad Técnica de Ambato, Ecuador.

FABIÁN MORALES, MSC, Universidad Técnica de Ambato, Ecuador.

DIEGO MORALES, MSC, Ministerio de Electricidad y Energías Renovables del Ecuador.

YOANDRYS MORALES TAMAYO, PHD, Universidad Técnica de Cotopaxi, Cotopaxi

OLENA LEONIDIVNA NAIDIUK, MSC, Universidad Politécnica Salesiana, Ecuador.

OSCAR NARANJO, MSC, Universidad del Azuay, Ecuador.

PAUL NARVAEZ, MSC, Universidad Politécnica Salesiana, Ecuador.

HERNÁN NAVAS OLMEDO, MSC, Universidad Técnica de Cotopaxi, Ecuador.

CESAR NIETO, PHD, Universidad Pontificia Bolivariana, Medellín, Colombia

FABIO OBANDO, MSC, Universidad Politécnica Salesiana, Ecuador.

LUIS ORTIZ FERNANDEZ, MSC, Universidade Federal de Rio Grande del Norte, Brasil

PABLO PARRA, MSC, Universidad Politécnica Salesiana, Ecuador.

PAULO PEÑA TORO, PHD, Ministerio de Productividad, Ecuador.

PATSY PRIETO VELEZ, MSC, Universidad Politécnica Salesiana, Ecuador.

DIEGO QUINDE FALCONI, MSC, Universidad Politécnica Salesiana, Ecuador.

DIANA QUINTANA ESPINOZA, MSC, Universidad Politécnica Salesiana, Ecuador.

WILLIAM QUITIAQUEZ SARZOSA, MSC, Universidad Politécnica Salesiana, Ecuador.

FLAVIO QUIZHPI PALOMEQUE, MSC, Universidad Politécnica Salesiana, Ecuador.

WASHINGTON RAMIREZ MONTALVAN, MSC, Universidad Politécnica Salesiana, Ecuador.

FRAN REINOSO AVECILLAS, MSC, Universidad Politécnica Salesiana, Ecuador.

NÉSTOR RIVERA CAMPOVERDE, MSC, Universidad Politécnica Salesiana, Ecuador.

JORGE ROMERO CONTRERAS, MSC, Universidad de Carabobo, Venezuela

FABIAN SAENZ ENDERICA, MSC, Universidad de las Fuerzas Armadas, ESPE, Ecuador.

LUISA SALAZAR GIL, PHD, Universidad Simón Bolívar, Venezuela

GUSTAVO SALGADO ENRÍQUEZ, MSC, Universidad Central del Ecuador., Ecuador.

JUAN CARLOS SANTILLÁN LIMA, MSC, Universidad Nacional de Chimborazo

JONNATHAN SANTOS BENÍTEZ, MSC, Universidad Politécnica Salesiana, Ecuador.

ANDRÉS SARMIENTO CAJAMARCA, MSC, Universidad Federal de Santa Catarina, Brasil

LUIS SERPA ANDRADE, MSC, Universidad Politécnica Salesiana, Ecuador.

CRISTIAN TIMBI SISALIMA, MSC, Universidad Politécnica Salesiana, Ecuador.

MILTON TIPAN SIMBAÑA, MSC, Universidad Politécnica Salesiana, Ecuador.

PAUL TORRES JARA, MSC, Universidad Politécnica Salesiana, Ecuador.

RODRIGO TUFÍÑO CÁRDENAS, MSC, Universidad Politécnica Salesiana, Ecuador.

FERNANDO URGILES ORTÍZ, MSC, Universidad Politécnica Salesiana, Ecuador.

JUAN VALLADOLID QUITOISACA, MSC, Universidad Politécnica Salesiana, Ecuador.

EFRÉN VÁZQUEZ SILVA, PHD, Universidad Politécnica Salesiana, Ecuador.

JULIO VERDUGO, MSC, Universidad Politécnica Salesiana, Ecuador.

MARY VERGARA PAREDES, PHD, Universidad de los Andes, Merida, Venezuela

JENNIFER YEPEZ ALULEMA, MSC, Universidad Politécnica Salesiana, Ecuador.

JULIO ZAMBRANO ABAD, MSC, Universidad Politécnica Salesiana, Ecuador.

PATRICIA ZAPATA MOLINA, MSC, Universidad Politécnica Salesiana, Ecuador.

Publications board

JUAN CÁRDENAS TAPIA, SDB, PHD
ESTEBAN MAURICIO INGA ORTEGA, PHD
ANGLE TORRES TOUKOUMIDIS, PHD
JAIME PADILLA VERDUGO, PHD
SHEILA SERRANO VINCENTI, MSC
JORGE CUEVA ESTRADA, MSC
JOHN CALLE SIGUENCIA, PHD
FLORALBA AGUILAR GORDÓN, PHD
BETTY RODAS SOTO, MSC
MÓNICA RUIZ VÁSQUEZ, MSC
JORGE ALTAMIRANO SÁNCHEZ, MSC
DAVID ARMENDÁRIZ GONZÁLEZ, MSC
JOSÉ JUNCOSA BLASCO, PHD

General Editor

ANGEL TORRES TOUKOUMIDIS, PHD

Technical board

DRA. MARCIA PEÑA, Style Reviewer,
Centro Gráfico Salesiano - Editorial Don Bosco
MSC. MARLON QUINDE ABRIL, Diagramming and layout
BSC. ANDRES LOPEZ, Community Manager - Diagramming and layout
BSC. MARÍA JOSÉ CABRERA, Marcalyc Support
BSC. CHRISTIAN ARPI, Community Managers Coordinator's team

Publications Service

HERNÁN HERMOSA (General Coordination)
MARCO GUTIÉRREZ (OJS Layout)
PAULINA TORRES (Style Editing)
RAYSA ANDRADE (Layout)
MARTHA VINUEZA (Layout)
YIXI GONZALEZ, (Style Reviewer)

Editorial

Editorial Abya Yala (Quito-Ecuador),
Av. 12 de octubre N422 y Wilson,
Bloque A, UPS Quito, Ecuador.
Casilla 17-12-719 Teléfonos: (593-2) 3962800 ext. 2638
email: editorial@abyayala.org

Printing: 800 copies

Typographic system used in the composition of this document \LaTeX .

INGENIUS

REVISTA DE CIENCIA Y TECNOLOGIA

Issue 33

january – june 2025

ISSN impreso 1390-650X / ISSN electrónico 1390-860X

The administration of the journal is done through the following parameters:

The journal uses the academic anti-plagiarism system  

The articles have an identification code (Digital Object Identifier) 

The editorial process is managed through the Open Journal System 

It is an open access publication (Open Access) licensed Creative Commons



The politics copyright of use postprint, are published in the Self-Archive Policy Repository

Sherpa/Romeo. 

The articles of the present edition can be consulted in

<http://revistas.ups.edu.ec/index.php/ingenius>



UNIVERSIDAD POLITÉCNICA SALESIANA DEL ECUADOR

INGENIUS Journal, is indexed in the following Databases and scientific information systems:

SELECTIVE DATABASES



Scopus



Google
scholar

SciELO Ecuador

BASE

EBCOhost
Fuente Académica Plus

Scientific Indexing Services

EuroPub

REVIEWS EVALUATION PLATFORMS

MIAR

Quality Open Access Market



SELECTIVE DIRECTORIES

latindex

Journal Seeker
Research Bible

AcademicKeys
UNLOCKING ACADEMIC CAREERS

SCIENCE LIBRARY INDEX

ULRICHSWEB™
GLOBAL SERIALS DIRECTORY

SELECTIVE SERIAL LIBRARY

redalyc.org

REDIB
Red Iberoamericana
de Innovación y Conocimiento Científico

refseek™

SCIENTIFIC LITERATURE SEARCHERS OPEN ACCESS

DOAJ
DIRECTORY OF
OPEN ACCESS
JOURNALS

EXALEAD

WORLDWIDE SCIENCE.ORG
The World's Open Access Journal

Journals for Free

OTHER BIBLIOGRAPHICAL DATABASES

Dialnet

Journal
TOCS
The latest Journal Tables of Contents

PKP|INDEX

CATALOG OF INTERNATIONAL UNIVERSITY LIBRARIES



Dear readers,

It is an honor to present this edition of our journal, showcasing the effort and talent of researchers from various countries who contribute to the advancement of science and technology in diverse contexts. The articles featured in this issue highlight the wide range of topics and approaches shaping technological and scientific development worldwide. From optimizing communication systems to innovations in air quality analysis and sustainable energy, this edition celebrates global collaboration and the local application of knowledge.

From Peru, researchers Abimael Adam Francisco Paredes, Heidy Velsy Rivera Vidal de Sánchez, Inés Eusebia Jesús Tolentino, and Jimmy Grover Flores Vidal present the article Determination of Optimal Formats for Digital Image Compression. This work analyzes methods to optimize the storage and transmission of images, a critical topic in the digital age.

Venezuela offers two remarkable contributions. In the first article, Monitoring the Liquid Level of a Coupled Tank System Using Quasi-LPV Control, Pedro Teppa-Garrán, Diego Muñoz-de Escalona, and Javier Zambrano explore solutions for efficient control in industrial systems. In the second, Comparative Analysis of Flow Patterns in Flat and Conical Nozzles Outside Design Conditions, San L. Tolentino and Jorge Mírez examine flow phenomena that impact the performance of mechanical devices.

Ecuador demonstrates its commitment to research and innovation through multiple contributions in this edition. Nancy Eras, José Andrés Otaño, and Santiago González present Application of MANETs as a Communication System for Sustainable Mobility, highlighting the impact of technology on urban sustainability.

The team of P. Silverio-Cevallos, J. Maita Cajarasca, D. A. Molina-Vidal, C. J. Tierra-Criollo, and P. Cevallos-Larrea introduces the Prototype

of a Multichannel Surface Muscle Stimulator Controlled Remotely, an innovative tool in the field of medicine.

In A Comprehensive Evaluation of AI Techniques for Predicting Air Quality Index: RNNs and Transformers, Pablo Andrés Buestán Andrade, Pedro Esteban Carrión Zamora, Anthony Eduardo Chamba Lara, and Juan Pablo Pazmiño Piedra explore the potential of artificial intelligence to address environmental challenges.

Ricardo Carpio-Chillo Gallo and Edwin Paccha-Herrera contribute an essential study titled Comparative Study of Cooling Strategies in a Lithium-Ion Battery Module to Prevent Thermal Runaway Using CFD, aimed at optimizing energy storage systems.

In the field of health, Darwin Patiño-Pérez, Luis Armijos-Valarezo, Luis Chóez-Acosta, and Freddy Burgos-Robalino investigate Convolutional Neural Networks for Diabetic Retinopathy Detection, addressing the impact of artificial intelligence in medical diagnosis.

Finally, Alan Cuenca Sánchez and Pablo Lluquiñga Eras present an educational approach in Design of a Didactic Energy Consumption Meter for Residential Use, demonstrating the importance of education in efficient energy use.

From China, Fengliang Qiao, Zhaojie Shen, and Yuxia Kang contribute the article Contact Patches of Radial Tires with Different Length-to-Width Ratios Under Static Load, providing practical solutions for the automotive industry and tire design.

The diversity of articles presented in this edition underscores the richness of international scientific talent. Contributions from Peru, Venezuela, Ecuador, and China bring fresh ideas and interdisciplinary approaches that inspire future collaborations and sustainable advancements. Each study, from its respective field, reaffirms the importance of science as a tool to address local challenges with global impact.

John Calle-Siguencia, PhD

Editor in Chief

TABLE OF CONTENTS

Determination of optimal formats for digital image compression	9
Determinación de los formatos óptimos para la compresión de imágenes digitales	
Abimael Adam Francisco Paredes, Heidy Velsy Rivera Vidal de Sánchez, Inés Eusebia Jesús Tolentino, Jimmy Grover Flores Vidal	
Liquid level tracking for a coupled tank system using quasi-lpv control	15
Seguimiento del nivel de líquido de un sistema de tanques acoplado empleando control cuasi-lpv	
Pedro Teppa-Garrán, Diego Muñoz-de Escalona, Javier Zambrano	
Application of MANETs as a communication system for sustainable mobility	27
Aplicación de MANET como sistema de comunicación en la movilidad sostenible	
Nancy Eras, José Andrés Otavalo, Santiago González	
Prototype of a remotely controlled multichannel surface muscle stimulator	38
Prototipo de estimulador muscular superficial multicanal controlado remotamente	
P. Silverio-Cevallos, J. Maita Cajamarca, D. A. Molina-Vidal, C. J. Tierra-Criollo, P. Cevallos-Larrea	
Contact patches of radial tires with different length-to-width ratios under static loads	49
Parches de contacto de neumáticos radiales con diferentes relaciones de longitud y anchura bajo carga estática	
Fengliang Qiao, Zhaojie Shen, Yuxia Kang	
A comprehensive evaluation of ai techniques for air quality index prediction: RNNs and transformers	60
Una evaluación integral de las técnicas de ia para predecir el índice de calidad del aire: RNN y transformers	
Pablo Andrés Buestán Andrade, Pedro Esteban Carrión Zamora, Anthony Eduardo Chamba Lara, Juan Pablo Pazmiño Piedra	
Comparative study of cooling strategies in a lithium-ion battery module for thermal runaway prevention using CFD	76
Estudio comparativo de estrategias de enfriamiento en un módulo de baterías de iones de litio para la prevención de la fuga térmica mediante CFD	
Ricardo Carpio-Chillo Gallo, Edwin Paccha-Herrera	
Convolutional neural networks for diabetic retinopathy detection	91
Redes neuronales convolucionales para detección de retinopatía diabética	
Darwin Patiño-Pérez, Luis Armijos-Valarezo, Luis Chóez-Acosta, Freddy Burgos-Robalino	
Design of a educational electrical energy consumption meter for residential use	102
Diseño de un medidor didáctico de consumo de energía eléctrica para uso residencial	
Alan Cuenca Sánchez, Pablo Llumiquinga Eras	
Comparative analysis of flow patterns in off-design planar and conical nozzle	115
Análisis comparativo de patrones de flujo en toberas planas y cónicas fuera de diseño	
San L. Tolentino, Jorge Mirez	
Normas editoriales	128
Guidelines	



DETERMINATION OF OPTIMAL FORMATS FOR DIGITAL IMAGE COMPRESSION

DETERMINACIÓN DE LOS FORMATOS ÓPTIMOS PARA LA COMPRESIÓN DE IMÁGENES DIGITALES

Abimael Adam Francisco Paredes^{1,*} , Heidy Velsy Rivera Vidal de Sánchez¹ ,
 Inés Eusebia Jesús Tolentino¹ , Jimmy Grover Flores Vidal¹ 

Received: 06-03-2024, Received after review: 21-06-2024, Accepted: 16-09-2024, Published: 01-01-2025

Abstract

The objective was to determine the influence of different image formats and tools used for compression on the final size of the images, to know which are the optimal formats for compression. The sample was made up of five digital image files with BMP extension, taken in different scenarios and at different times at the researcher's discretion. The technique used was the analysis of digital image files and as an instrument a double input matrix, where the conversions of BMP files to six different extensions of image files were registered, with four different tools for manipulation of image files. The experimental design was factorial, where the two factors were the image compression formats and tools and the dependent variable the final image file size. Factorial ANOVA statistical analysis was applied with $\alpha = 0.05$. It was obtained that the format of smaller size was the JPG when using as tool the Illustrator and the one of greater size the one of greater extension the PSD also obtained with the Illustrator. The statistical analysis showed that the format factor significantly influences the final size of the images ($p < 0.05$) and the tool factor does not show significant influence on the size of the images ($p > 0.05$), nor is the interaction between the factors significant. It is concluded that regardless of the tool used, it is the image format that influences the final size.

Keywords: image quality, compression techniques, pixels, image handlers, size reduction

Resumen

El objetivo de este trabajo fue el determinar la influencia de diferentes formatos de imagen y herramientas que se utilizan para la compresión en el tamaño final de las mismas, para conocer cuáles son los formatos óptimos para la compresión. La muestra estuvo conformada por cinco archivos de imágenes digitales con extensión .bmp, tomadas en diferentes escenarios y horas a criterio del investigador. La técnica empleada fue el análisis de archivos de imágenes digitales y como instrumento una matriz de doble entrada, donde se registraron las conversiones de los archivos .bmp a seis diferentes extensiones de archivos de imágenes, con cuatro diferentes herramientas de manipulación de archivos de imágenes. El diseño experimental fue factorial, donde los dos factores fueron los formatos y las herramientas de compresión de imágenes y la variable dependiente, el tamaño final del archivo de imagen. Se aplicó análisis estadístico ANOVA factorial con $\alpha = 0,05$. Se obtuvo que el formato de menor tamaño fue el .jpg al utilizar como herramienta el *Illustrator* y el de mayor tamaño el .psd, también obtenido con el *Illustrator*. El análisis estadístico mostró que el factor formato influye de forma significativa en el tamaño final de las imágenes ($p < 0,05$) y el factor herramienta no muestra influencia significativa en el tamaño de las imágenes ($p > 0,05$), como tampoco es significativa la interacción entre los factores. Se concluye que independientemente de la herramienta que se utilice, es el formato de la imagen lo que influye en el tamaño final.

Palabras clave: calidad de imagen, técnicas de compresión, píxeles, manejadores de imágenes, reducción de tamaño

^{1,*}Universidad Nacional Hermilio Valdizán, Pillco Marca, Perú.
 Corresponding author ✉: adamfp28@hotmail.com.

Suggested citation: Paredes, A.; Rivera Vidal de Sánchez, H. ; Tolentino, I. and Flores Vidal, J. "Determination of optimal formats for digital image compression," *Ingenius, Revista de Ciencia y Tecnología*, N.º 33, pp. 9-14, 2025, DOI: <https://doi.org/10.17163/ings.n33.2025.01>.

1. Introduction

Image compression encompasses a set of techniques applied to digital images that enable efficient storage or transmission [1]. These techniques have been developed to address the significant file sizes that image files can occupy, which often limits their exchange via email and other electronic platforms. Compression methods rely on mathematical algorithms that reduce file size, thereby minimizing resource consumption and transfer time [2].

All image compression algorithms aim to achieve a smaller compressed image size (high compression factor) while maintaining a high-quality reconstructed image (high-quality compression). The efficiency of these algorithms can be evaluated based on the specific application using various criteria [3]. The most important criterion is the compression factor, which compares the image size before and after compression. Therefore, a higher compression factor indicates a more effective compression algorithm [4].

The most common types of compression are lossless and lossy compression. In lossy compression, some image information is discarded during the compression process. Some algorithms combine both techniques to achieve compression [5]. The effectiveness of compression also depends on the type of image. For example, bitmap images are composed of a grid of cells or pixels, each with a specific size, and lose resolution when resized [2]. In contrast, vector images are constructed from mathematically defined objects, such as points and lines, which are controlled through Bézier curves. This structure provides greater flexibility, as vectors can be scaled without losing resolution [6].

Regarding compression, both with and without loss, Ruiz, Yarasca and Ruiz [7] explain that lossless compression employs complex mathematical algorithms to condense code chains while preserving all the information in the image. This ensures that the image can be fully regenerated, without any loss, during decompression, although it requires specific encoding and decoding times. Formats such as Portable Network Graphics (PNG) utilize this type of compression. In contrast, Rojatkari et al. [8] describe lossy compression as a technique in which certain image information, typically deemed minimally perceptible, is discarded during compression, resulting in a loss of some original file data. This method is commonly used in image formats such as Joint Photographic Experts Group (JPEG).

Among the most common types of lossless image compression is Run Length Encoding (RLE), which, as noted by Hardi et al. [9], is one of the simplest compression schemes. It works by replacing sequences of identical bits with a code. The method scans the image to identify pixels of the same color, and when the image is saved, only the color value and the position of the

color pixels are recorded. This technique is particularly effective for images with large areas of uniform color, as it compresses the image without losing quality. The Lempel-Ziv-Welch (LZW) method is similar to RLE but supports a wider range of formats, including TIFF, PDF, and GIF [2]. Like RLE, it is highly effective for images with large areas of uniform color and simple designs, but its efficiency diminishes when compressing images with a broad range of photographic-like colors. Huffman coding assigns shorter bit codes to frequently occurring data and longer codes to less frequent data, making it widely used due to its simplicity and high speed [1]. Arithmetic coding, conversely, represents sequences of symbols in binary form by using intervals of real values between zero and one [10].

The most common lossless compression models include Transform Coding, which uses a discrete Fourier transform to represent the image through transform coefficients. A quantization process is then applied, where coefficients with small, insignificant values are eliminated, resulting in some loss of information without causing noticeable image distortion [11]. Vector quantization involves selecting a representative set of pixels from the original image and discarding the non-representative ones. This is achieved by constructing dynamic tables or through clustering for vector classification [1]. Fractal Compression treats images as fractal objects, meaning they are composed of a repeating fragmented structure. Then, some functions are created to generate transformations that divide the original image into smaller, self-similar parts. The iterative application of these transformations produces an image that closely resembles the original but is smaller in size, as some information is lost during the division process [10].

The compression of digital images has been studied from various perspectives, ranging from research on the algorithms used for compression [12,13] [3,4] to specific applications in fields such as forestry sciences [14], forensic sciences [15,16], medical sciences [17–19], and other disciplines. Most research focuses on algorithms for analysing digital image files, with little reference to comparisons of tools and formats for selecting optimal options. This study aims to address this gap by examining whether image formats, tools, and their interaction influence the size of image files. The primary objective of the research is to determine whether the interaction between formats and tools significantly affects the size of digital image files.

2. Materials and methods

2.1. Methodology

The research followed a quantitative methodological approach, utilizing an experimental design at an explanatory level. The population consisted of photo-

graphic image files in BMP format, and the sample included five digital image files in BMP format, captured in various settings and at different times at the researcher's discretion. The technique employed was the analysis of digital image files, and the instrument used was a double-entry matrix, where the conversions of BMP files to six different file extensions were recorded using four different tools (software) for digital image manipulation.

The statistical design employed was a factorial design, with two factors: image compression formats and tools, and the dependent variable being the final image file size. The experimental design included six levels for the first factor (Format) and four levels for the second factor (Tool), resulting in a total of 24 treatments applied to a sample of five images, yielding 120 measurements of the dependent variable. The formats used were JPG, PNG, PSD, PDF, TIFF, and TGA, while the image editing tools were CorelDraw, Photoshop, Illustrator, and Gimp. The images selected for the sample were coded as presented in Table 1.

Table 1. Sample description

N°	Name	Extension
1	Archive f01_bmp.bmp	11796536 bytes
2	Archive f02_bmp.bmp	47775800 bytes
3	Archive f03_bmp.bmp	47775800 bytes
4	Archive f04_bmp.bmp	53747768 bytes
5	Archive f05_bmp.bmp	59722040 bytes

The statistical analysis was conducted using factorial ANOVA to measure both the individual effects of each factor and the effect of their interaction on the dependent variable, with a 95% confidence level. The statistical software used for the analysis was SPSS 25.

3. Results and Discussion

3.1. Analysis and Results

Of the 120 data points recorded, it was observed that when using the CorelDraw tool, the PDF format produced the smallest files [20], while the TGA format resulted in the largest. With the Photoshop tool, the smallest file size was achieved using the JPG format, while the largest was obtained with TIFF. Similarly, when using Illustrator, the JPG format yielded the smallest file size, whereas the PSD format generated the largest files. Lastly, when using Gimp, the smallest file size was also achieved with the JPG format, while the PSD format resulted in the largest files. As noted, JPG consistently produced the smallest file sizes. Although JPG compression is lossy, the algorithm compensates by softening edges and areas with similar colors, making the loss of information imperceptible to the naked eye. This allows for a high degree

of compression, with image quality degradation only noticeable under significant zoom [2]. In this regard, Tan [19] suggests that the choice of format should primarily be based on the content of the image. Photographic images, or those with soft tones and few sharp edges, are generally best compressed using a lossy format such as JPEG.

The analysis also revealed that, across the four tools used for converting BMP files, the conversion to JPG resulted in the lowest average file size, at 210,152,480 bytes. In contrast, the conversion to PSD produced the highest average file size, at 7,042,890,180 bytes. The smallest file was a JPG generated with the Illustrator tool, with a size of 50,474,220 bytes, while the largest file was a PSD, also generated with Illustrator, with a size of 10,044,601,140 bytes. As noted on the Adobe Photoshop portal, PSD is the default format for Photoshop and is compatible with other tools, such as Illustrator. PSD files can reach a maximum size of 2 GB. The fact that the average PSD file size in this study exceeded that limit suggests that some tools may not be fully optimized for converting BMP images to PSD format. Illustrator, which produced PSD files of approximately 1 GB, proved to be the most effective for this conversion. Additionally, Parmar and Pancholi [21] mention that the JPG format is widely used in photography due to its ability to handle millions of colors while maintaining good quality, even with lossy compression.

A factorial ANOVA test was performed to assess the influence of the factors (Format and Tool) on the dependent variable (Image size), with the results presented in Table 2.

Table 2. Results of the applied factorial ANOVA

Origen	Sum of squares	gl	Mean square	P
Model	8.52022E+20	23	3.70444E+19	0
Intersection	1.00536E+21	1	1.00536E+21	0
A. Formats	5.63869E+20	5	1.12774E+20	0
B. tools	5.86365E+19	3	1.95455E+19	0.09
A*B	2.29516E+20	15	1.53011E+19	0.053
Error	8.4412E+20	96	8.79292E+18	
Total	2.7015E+21	120		
Total	1.69614E+21	119		

In Table 2, the p-value is of primary importance, as it indicates the influence of the factors on the dependent variable. For the Format factor, $p < 0.05$ suggests a statistically significant influence on the size of the final converted file, which is expected given the variation in file sizes across different formats. In contrast, the Tool factor shows no significant influence, as $p > 0.05$, indicating that the final image size is not dependent on the tool used for conversion. Additionally, the interaction between the factors does not exhibit any significant influence on the final image size. Therefore,

it can be concluded that the format is the primary factor affecting the final size of BMP images, regardless of the tool used for conversion.

The differences between format types are expected and consistent with the fact that each format employs a distinct algorithm for compression, which directly influences the final file size [1]. This observation is also supported by Salomón [22], who highlights that each format uses a different compression methodology, and, therefore, conversion and compression tools, when operating based on these methodologies, do not show significant differences between them. The results indicate that compressing an image to a particular format can be achieved using any tool, as the resulting file size will not be statistically different. This is further supported by the non-significance of the interaction between the factors. Similarly, AbuBaker, Eshtay, and AkhoZahia [12] reported differences in the size and quality of digital mammogram images depending on the compression methods used, a trend also evident in the differences between the various output formats, each utilizing distinct methods. Likewise, Wahba and Maghari [23] demonstrated that the compression techniques unique to each format are key determinants of the file size or extension of the compressed image.

When differences between formats were observed, JPG consistently resulted in the smallest file sizes while maintaining acceptable quality, comparable to other formats. This finding aligns with Dhawan's [24] research, which compared the compression of various image formats based on the different algorithms used. The smallest JPG file sizes were obtained using the Illustrator tool, suggesting that, although no statistical difference was found between the tools, Illustrator may be preferable for achieving smaller file sizes. This is further supported by Sakshica and Gupta [25,26], who emphasize that Illustrator is particularly effective for compressing vector images [27–30].

4. Conclusions

The results of this study indicate that the final file size of compressed images is determined primarily by the format chosen for compression rather than the tool employed. The smallest file size was consistently achieved with the JPG format, particularly when using the Illustrator tool, which is notably effective for compressing vector images.

However, while JPG yielded the smallest file size, it employs a lossy compression method, which results in the loss of some image pixels, potentially affecting resolution upon decompression. Therefore, further research and experimentation with alternative tools are recommended to more effectively determine the optimal image format for compression, ensuring a balance between minimal file size and the preservation of

image quality aligned with the intended use.

References

- [1] N. La Serna, L. Pro Concepción, and C. Yañez Durán, “Compresión de imágenes: Fundamentos, técnicas y formatos,” *Revista de Ingeniería de Sistemas e Informática*, vol. 6, no. 1, pp. 21–29, 2009. [Online]. Available: <https://upsalesiana.ec/ing32ar1r01>
- [2] C. A. Ordoñez Santiago, “Formatos de imagen digital,” *Revista Digital Universitaria*, vol. 5, no. 7, 2005. [Online]. Available: <https://upsalesiana.ec/ing32ar1r02>
- [3] M. Al-khassaweneh and O. AlShorman, “Frei-chen bases based lossy digital image compression technique,” *Applied Computing and Informatics*, vol. 20, no. 1/2, pp. 105–118, 2024. [Online]. Available: <https://doi.org/10.1016/j.aci.2019.12.004>
- [4] AlShorman, O. M. Mahmoud, AlKhassaweneh, and Mahmood, “Lossy digital image compression technique using run-length encoding and frei-chen basis,” in *Universidad de Yarmouk*, 2012. [Online]. Available: <https://upsalesiana.ec/ing32ar1r4>
- [5] P. Chamorro-Posada, “A simple method for estimating the fractal dimension from digital images: The compression dimension,” *Chaos, Solitons & Fractals*, vol. 91, pp. 562–572, 2016. [Online]. Available: <https://doi.org/10.1016/j.chaos.2016.08.002>
- [6] L. Arranz, *Vector images and bitmaps*. Recursostic, 2005. [Online]. Available: <https://upsalesiana.ec/ing32ar1r6>
- [7] M. E. Ruiz Rivera and E. Yarasca Carranza, Juan Eduardo Ruiz Lizama, “Análisis de la compresión de imágenes utilizando clustering bajo el enfoque de colonia de hormigas,” *Industrial Data*, vol. 16, no. 2, pp. 118–131, 2013. [Online]. Available: <https://doi.org/10.15381/idata.v16i2.11929>
- [8] D. V. Rojatkhar, N. D. Borkar, B. R. Naik, and R. N. Peddiwar, “Image compression techniques: Lossy and lossless,” in *International Journal of Engineering Research and General Science*, vol. 3, no. 2, 2015, pp. 912–917. [Online]. Available: <https://upsalesiana.ec/ing32ar1r66>
- [9] S. M. Hardi, B. Angga, M. S. Lydia, I. Jaya, and J. T. Tarigan, “Comparative analysis run-length encoding algorithm and fibonacci code algorithm on image compression,” *Journal of Physics: Conference Series*, vol. 1235, no. 1, p. 012107, jun 2019. [Online]. Available: <https://dx.doi.org/10.1088/1742-6596/1235/1/012107>

- [10] G. E. Belloch, *Introduction to Data Compression*. Computer Science Department. Carnegie Mellon University, 2013. [Online]. Available: <https://upsalesiana.ec/ing32ar1r10>
- [11] R. C. González and R. E. Woods, *Tratamiento digital de imágenes*. Madrid: Díaz de Santos, 1996. [Online]. Available: <https://upsalesiana.ec/ing32ar1r11>
- [12] A. AbuBaker, M. Eshtay, and M. AkhoZahia, “Comparison study of different lossy compression techniques applied on digital mammogram images,” *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 12, pp. 149–155, 2016. [Online]. Available: <http://dx.doi.org/10.14569/IJACSA.2016.071220>
- [13] C. Ding, Y. Chen, Z. Liu, and T. Liu, “Implementation of grey image compression algorithm based on variation partial differential equation,” *Alexandria Engineering Journal*, vol. 59, no. 4, pp. 2705–2712, 2020, new trends of numerical and analytical methods for engineering problems. [Online]. Available: <https://doi.org/10.1016/j.aej.2020.05.012>
- [14] X. P. Alaitz Zabala, R. Díaz-Delgado, F. García, F. Auli-Llinas, and J. Serra-Sagrasta, “Effects of jpeg and jpeg2000 lossy compression on remote sensing image classification for mapping crops and forest areas,” *e Ministry of Science and Technology and the FEDER*, 2020. [Online]. Available: <https://upsalesiana.ec/ing32ar1r14>
- [15] M. C. Stamm and K. J. R. Liu, “Anti-forensics of digital image compression,” *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 1050–1065, 2011. [Online]. Available: <http://doi.org/10.1109/TIFS.2011.2119314>
- [16] T. H. Thai, R. Cogramne, F. Retraint, and T.-N.-C. Doan, “Jpeg quantization step estimation and its applications to digital image forensics,” *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 1, pp. 123–133, 2017. [Online]. Available: <http://doi.org/10.1109/TIFS.2016.2604208>
- [17] L. González, J. Muro, M. del Fresno, and R. Barbuza, *Un enfoque para la compresión de imágenes médicas basado en regiones de interés y compensación de movimiento*. 4to Congreso Argentino de Informatica y Salud, CAIS 2013, 2013. [Online]. Available: <https://upsalesiana.ec/ing32ar1r17>
- [18] F. Liu, M. Hernandez-Cabronero, V. Sanchez, M. W. Marcellin, and A. Bilgin, “The current role of image compression standards in medical imaging,” *Information*, vol. 8, no. 4, 2017. [Online]. Available: <https://doi.org/10.3390/info8040131>
- [19] M. A. Ameer Kadhum, “Compression the medical images using length coding method,” *Journal of Electrical and Electronics Engineering*, vol. 12, no. 3, pp. 94–98, 2017. [Online]. Available: <http://doi.org/10.9790/1676-1203029498>
- [20] Adobe. (2023) Elección de un formato de archivo. Adobe. All rights reserved. [Online]. Available: <https://upsalesiana.ec/ing32ar1r26>
- [21] C. K. Parmar and K. Pancholi, “A review on image compression techniques,” *Journal of Information, Knowledge and Research in Electrical Engineering*, vol. 2, no. 2, pp. 281–284, 2013. [Online]. Available: <https://upsalesiana.ec/ing32ar1r20>
- [22] D. Salomon, G. Motta, and D. Bryant, *Compresión de datos. La referencia completa*. Springer-Verlag London Limited, 2007. [Online]. Available: <https://upsalesiana.ec/ing32ar1r21>
- [23] W. Wahba and A. Maghari, “Lossless image compression techniques comparative study,” *International Research Journal of Engineering and Technology (IRJET)*, vol. 3, 02 2016. [Online]. Available: <https://upsalesiana.ec/ing32ar1r22>
- [24] S. Dhawan, “A review of image compression and comparison of its algorithms,” *International Journal of Electronics & Communication Technology*, vol. 2, no. 1, pp. 22–26, 2011. [Online]. Available: <https://upsalesiana.ec/ing32ar1r23>
- [25] K. Sakshica and K. Gupta, “Various raster and vector image file formats,” *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 4, no. 3, pp. 268–271, 2015. [Online]. Available: <http://doi.org/10.17148/IJARCCE.2015.4364>
- [26] A. K. Al-Janabi, “Efficient and simple scalable image compression algorithms,” *Ain Shams Engineering Journal*, vol. 10, no. 3, pp. 463–470, 2019. [Online]. Available: <https://doi.org/10.1016/j.asej.2019.01.008>
- [27] V. Barannik, S. Sidchenko, N. Barannik, and V. Barannik, “Development of the method for encoding service data in cryptocompression image representation systems,” *Eastern-European Journal of Enterprise Technologies*, vol. 3, no. 9, pp. 103–115, 2021. [Online]. Available: <https://doi.org/10.15587/1729-4061.2021.235521>
- [28] P. K. Pareek, C. Sridhar, R. Kalidoss, M. Aslam, M. Maheshwari, P. K. Shukla, and S. J. Nuagah,

- “Intopmicm: Intelligent medical image size reduction model,” *Journal of Healthcare Engineering*, vol. 2022, no. 1, p. 5171016, 2022. [Online]. Available: <https://doi.org/10.1155/2022/5171016>
- [29] X. Gao, J. Mou, S. Banerjee, and Y. Zhang, “Color-gray multi-image hybrid compression–encryption scheme based on bp neural network and knight tour,” *IEEE Transactions on Cybernetics*, vol. 53, no. 8, pp. 5037–5047, 2023. [Online]. Available: <https://doi.org/10.1109/TCYB.2023.3267785>
- [30] R. Kumar, P. Seetharaman, A. Luebs, I. Kumar, and K. Kumar, “High-fidelity audio compression with improved rvqgan,” *Advances in Neural Information Processing Systems*, 2023. [Online]. Available: <https://doi.org/10.48550/arXiv.2306.06546>



LIQUID LEVEL TRACKING FOR A COUPLED TANK SYSTEM USING QUASI-LPV CONTROL

SEGUIMIENTO DEL NIVEL DE LÍQUIDO DE UN SISTEMA DE TANQUES ACOPLADO EMPLEANDO CONTROL CUASI-LPV

Pedro Teppa-Garrán^{1,*} , Diego Muñoz-de Escalona¹ ,
 Javier Zambrano¹ 

Received: 13-03-2024, Received after review: 21-05-2024, Accepted: 03-07-2024, Published: 01-01-2025

Abstract

This article proposes a gain-scheduling procedure based on quasi-LPV modeling for a nonlinear coupled tank system to track the liquid level with zero steady-state error. The nonlinearities are directly represented by a parameter vector that varies within a bounded set constrained by the physical limits of the tank system levels. This approach enables accurate nonlinear system modeling using a linear parameter-varying model. State-feedback linear controllers are designed at the extreme vertices of the bounded set. The global controller is derived as the weighted average of local controller contributions, with the weighting determined by the instantaneous values of the parameter vector. Two interpolation mechanisms are proposed to implement this weighted averaging of the linear controllers. The results confirm the effectiveness of the proposed method in achieving accurate liquid level tracking.

Keywords: Coupled-Tank System, Gain Scheduling, Nonlinear Systems, Quasi-LPV, Tracking Problem

Resumen

En este artículo se propone un procedimiento de programación de ganancias basado en un modelado cuasi-LPV de un sistema no lineal de tanques acoplados para seguir el nivel de líquido con error en estado estacionario nulo. Las no linealidades están representadas directamente por un vector de parámetros que varía dentro de un conjunto acotado por los límites físicos del nivel del sistema de tanques. Esto permite un modelado exacto del sistema no lineal utilizando un modelo lineal de parámetros variantes. Luego, se diseñan controladores lineales de realimentación de estado en los vértices extremos del conjunto acotado. El controlador global corresponde a un promedio ponderado de las contribuciones locales. Esta ponderación depende de los valores instantáneos del vector de parámetros. Para implementar el promedio ponderado de los controladores lineales, se proponen dos mecanismos de interpolación. Los resultados obtenidos muestran la efectividad del método.

Palabras clave: sistema de tanques acoplados, programación de ganancias, sistemas no lineales, cuasi-LPV, problema de seguimiento

^{1,*}Departamento de Gestión de Proyectos y Sistemas, Universidad Metropolitana, Caracas, Venezuela.
 Corresponding author ✉: pteppa@unimet.edu.ve.

Suggested citation: Teppa-Garrán, P.; Muñoz-de Escalona, D. and Zambrano, J. "Liquid level tracking for a coupled tank system using quasi-lpv control," *Ingenius, Revista de Ciencia y Tecnología*, N.º 33, pp. 15-26, 2025, DOI: <https://doi.org/10.17163/ings.n33.2025.02>.

1. Introduction

The control of liquid levels in tanks is widely employed in various industries, including food and beverage production, nuclear and petrochemical plants, and the pharmaceutical sector. Generally, interactions between tanks occur due to coupling, resulting in nonlinear behavior [1]. Numerous control strategies have been proposed for coupled tank systems, including Proportional-Integral-Derivative (PID) controllers [2–4], Fuzzy control [5,6], Model Predictive Control [7,8], Backstepping Control [9,10], Sliding-Mode Control [11,12], Fractional PID controllers [13,14], Robust control [15]. Active Disturbance Rejection Control [16,17] and Two-Degree-Of-Freedom controllers [18]. Some of these techniques rely on nonlinear system theory, which can be challenging to implement, while others employ linearization of the system equations around an operating point. For the local operating range, designs based on Jacobian linearization perform effectively. However, under significant disturbances or when faster settling times are required, the performance of such controllers can deteriorate due to a loss of robustness.

Gain scheduling [19,20] is a widely adopted approach in industry for controlling nonlinear systems by breaking down the nonlinear design problem into several smaller, manageable subproblems where linear design tools can be applied. For instance, in robot control, controller dynamics are adjusted in real-time based on varying inertia and geometry. Similarly, most aircraft control laws are modified by interpolating individually designed controllers. In recent decades, the Linear Parameter Varying (LPV) system theory has gained prominence as a powerful paradigm for system identification, analysis, and controller synthesis [21–23]. This class of systems is particularly valuable as it allows nonlinearities to be incorporated as varying parameters within a bounded set, ensuring that the possible trajectories of the LPV system encompass all trajectories of the original nonlinear system. When these parameters include state vector elements, the system is referred to as quasi-LPV [24]. In this study, the nonlinearities of the tank system model, represented by liquid levels, are considered uncertain but constrained within the technological limits of the equipment ($[0, 30]cm$). This allows for an accurate representation of the nonlinear terms by embedding them into a quasi-LPV model. The advantage of this approach is that it enables the design of linear controllers using state-space techniques, ensuring zero steady-state tracking error for constant reference inputs and guaranteeing a pole-dominant criterion [25,26].

Within a gain-scheduling scheme, the control of the nonlinear coupled tank system is achieved through local controller interpolation. Two interpolation mechanisms are proposed: (1) analytical interpolation, where a system of linear equations is continuously solved to

compute the weighting factors, and (2) geometric interpolation, where the weights of the local controllers are determined based on the Euclidean distance to some vertex points. Analytical interpolation, initially presented in [27] and inspired by concepts from Takagi-Sugeno fuzzy models, is implemented in this study in a simplified form without incorporating any fuzzy elements. Geometric interpolation, on the other hand, offers an innovative approach in this context. While quasi-LPV theory has been widely applied in fields such as missile guidance [28,29] and robotics [30,31], its application to tank systems remains relatively unexplored despite the significant industrial relevance of this process.

The results confirm the effectiveness of the proposed method in controlling the coupled tank system. The article is structured as follows: Section 2 details the quasi-LPV control design method, with a particular focus on the formulation of two interpolation mechanisms, which are integrated within a gain-scheduled tracking control strategy and describes the coupled tank system's nonlinear model. In Section 3, the quasi-LPV design method is applied to the system. Finally, the conclusions are presented in Section 4.

Notación: Bold capital letters denote matrices, while bold lowercase letters represent vectors (i.e. θ_j is the j -th component of the vector $\boldsymbol{\theta}$). Superscripts indicate vectors; for instance, $\boldsymbol{\theta}^i$ refers to the i -th vector, and θ_j^i denotes the j -th component of the i -th vector. $\dot{f}(t) = df(t)/dt$ and $\ddot{f}(t) = d^2f(t)/dt^2$. \mathbb{R} denotes the set of real numbers.

2. Materials and methods

2.1. Quasi-LPV control design

Most existing nonlinear controller synthesis approaches focus on input-affine systems [32], which are typically described as equation (1):

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}[\boldsymbol{x}(t)] + \boldsymbol{g}[\boldsymbol{x}(t)]u(t) \quad (1)$$

Where $\boldsymbol{x} : \mathbb{R}_+ \rightarrow \mathbb{R}^n$ is the state vector, $u : \mathbb{R}_+ \rightarrow \mathbb{R}$ is the control input, t is the independent variable of time, $\boldsymbol{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $\boldsymbol{g} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ are nonlinear fields. The general nonlinear equation $\dot{\boldsymbol{x}}(t) = \boldsymbol{f}[\boldsymbol{x}(t), u(t)]$, which frequently appears, can, under technical assumptions, be transformed into (1) through a nonlinear feedback transformation [33].

The first step in the synthesis procedure is to derive a quasi-LPV representation of the form as seen in the equation (2):

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}[\boldsymbol{\theta}(t)]\boldsymbol{x}(t) + \boldsymbol{B}[\boldsymbol{\theta}(t)]u(t), \boldsymbol{\theta} \in \Omega \quad (2)$$

For the nonlinear system described in equation (1). Here $\boldsymbol{\theta}$ represents a parameter vector that varies

within a bounding box Ω . In quasi-LPV modeling, it is assumed that there is a relationship between the parameter vector and the system states, $\boldsymbol{\theta} = \boldsymbol{\sigma}(\mathbf{x})$, such that for all parameter values within Ω .

$$\mathbf{A}[\boldsymbol{\theta}(t)]\mathbf{x}(t) + \mathbf{B}[\boldsymbol{\theta}(t)]u(t) = \mathbf{f}[\mathbf{x}(t)] + \mathbf{g}[\mathbf{x}(t)]u(t)$$

2.2. Interpolation mechanisms

Defining $\boldsymbol{\theta}^i \in \mathbb{R}^n$, $i = 1, \dots, N$ in equation (2) as the vectors representing the extreme combinations of parameters in Ω , a set of local linearized models is obtained as follows:

$$(\mathbf{A}_i, \mathbf{B}_i) = (\mathbf{A}(\boldsymbol{\theta}^i), \mathbf{B}(\boldsymbol{\theta}^i)), i = 1, \dots, N \quad (3)$$

For each local model, a state vector gain \mathbf{K}_i can be designed. The parameter vector $\boldsymbol{\theta}(t)$ is then used to construct the overall gain-scheduled controller by interpolating the local controllers. At any given time, $\boldsymbol{\theta}(t)$ can be expressed as equation (4):

$$\boldsymbol{\theta}(t) = \sum_{i=1}^N \alpha_i(t) \boldsymbol{\theta}^i, \quad \sum_{i=1}^N \alpha_i(t) = 1 \quad (4)$$

The weights $\alpha_i(t)$ are computed by solving the system of linear equations:

$$\mathbf{W}\boldsymbol{\eta} = \mathbf{v} \quad (5)$$

Where:

$$\boldsymbol{\eta} = \begin{bmatrix} \alpha_1(t) \\ \vdots \\ \alpha_N(t) \end{bmatrix}, \quad \mathbf{W} = \begin{bmatrix} \boldsymbol{\theta}^1 & \dots & \boldsymbol{\theta}^N \\ 1 & \dots & 1 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} \boldsymbol{\theta}(t) \\ 1 \end{bmatrix}$$

The interpolation scheme based on the weights computed from the continuous solution of equation (5) is referred to as analytical interpolation to distinguish it from geometric interpolation, which will be described below.

At any given time, the Euclidean distance between the state-dependent parameter vector $\boldsymbol{\theta} \in \mathbb{R}^n$ and any of the extreme vectors $\boldsymbol{\theta}^i \in \mathbb{R}^n$ within the bounding box Ω can be computed for $i = 1, \dots, N$ as:

$$\overline{\boldsymbol{\theta}\boldsymbol{\theta}^i} = \sqrt{(\theta_1 - \theta_1^i)^2 + \dots + (\theta_n - \theta_n^i)^2} \quad (6)$$

The relative contribution of the parameter vector $\boldsymbol{\theta}(t)$ on each vertex $\boldsymbol{\theta}^i$ is given by:

$$a_j = \frac{\overline{\boldsymbol{\theta}\boldsymbol{\theta}^j}}{\sum_i \overline{\boldsymbol{\theta}\boldsymbol{\theta}^i}}, j = 1, \dots, N$$

Points further from the vertices should have lower weights. Therefore, the complementary distance, $1 -$

$\overline{\boldsymbol{\theta}\boldsymbol{\theta}^i}$, is used in the calculation. The complementary relative contribution is then computed as $b_j = 1 - a_j$, $j = 1, \dots, N$. Finally, the weight $\alpha_j(t)$ para $j = 1, \dots, N$ at any given time is determined as:

$$\alpha_j = \frac{b_j}{\sum_{i=1}^N b_i} = \frac{1 - \frac{\overline{\boldsymbol{\theta}\boldsymbol{\theta}^j}}{\sum_{i=1}^N \overline{\boldsymbol{\theta}\boldsymbol{\theta}^i}}}{\sum_{j=1}^N \left(1 - \frac{\overline{\boldsymbol{\theta}\boldsymbol{\theta}^j}}{\sum_{i=1}^N \overline{\boldsymbol{\theta}\boldsymbol{\theta}^i}}\right)} \quad (7)$$

As in the analytical procedure, the weights computed using the geometric approach continuously satisfy the equation $\sum_{j=1}^N \alpha_j(t) = 1$. The key difference between the two methods is that the geometric approach ensures positive weights, whereas the analytical procedure does not. This may necessitate conditioning of the control input if the actuator operates only with positive signals.

2.3. Tracking a step reference input

Using the computed weights, the model in equation (2) can be approximated as a combination of the local linear models:

$$\dot{\mathbf{x}}(t) = \underbrace{\sum_{i=1}^N \alpha_i(t) \mathbf{A}_i}_{\mathbf{A}} \mathbf{x}(t) + \underbrace{\sum_{i=1}^N \alpha_i(t) \mathbf{B}_i}_{\mathbf{B}} \mathbf{u}(t) \quad (8)$$

The design problem now focuses on tracking a step reference input $r(t)$ with zero steady-state error $e(t)$ defined as:

$$e(t) = r(t) - y(t) \quad (9)$$

Where is the controlled output. Taking the time derivative of equation (9), for a constant reference input yields:

$$y(t) = \mathbf{C}\mathbf{x}(t) \quad (10)$$

$$\dot{e}(t) = -\mathbf{C}\dot{\mathbf{x}}(t) \quad (11)$$

Taking the time derivative of each local linear model $(\mathbf{A}_i, \mathbf{B}_i)$ for $i = 1, \dots, N$ yields:

$$\dot{\mathbf{x}}(t) = \mathbf{A}_i \dot{\mathbf{x}}(t) + \mathbf{B}_i \dot{\mathbf{u}}(t) \quad (12)$$

Equations (11) and (12) can be combined as:

$$\dot{\mathbf{z}}(t) = \mathbf{F}_i \mathbf{z}(t) + \mathbf{G}_i u_0(t) \quad (13)$$

Where:

$$\mathbf{z}(t) = [e(t) \quad \dot{\mathbf{x}}(t)]^T, \quad u_0(t) = \dot{\mathbf{u}}(t),$$

$$\mathbf{F}_i = \begin{bmatrix} 0 & -\mathbf{C} \\ \mathbf{0} & \mathbf{A}_i \end{bmatrix}, \quad \mathbf{G}_i = \begin{bmatrix} 0 \\ \mathbf{B}_i \end{bmatrix}.$$

A state feedback gain for system (13) is constructed as:

$$u_0(t) = \mathbf{K}_i \mathbf{z}(t) = \begin{bmatrix} K_{e_i} & \mathbf{K}_{x_i} \end{bmatrix} \begin{bmatrix} e(t) \\ \dot{\mathbf{x}}(t) \end{bmatrix} \quad (14)$$

After integrating equation (14), the actual control signal becomes:

$$u(t) = K_{e_i} \int_0^t e(\tau) d\tau + \mathbf{K}_{x_i} \mathbf{x}(t) \quad (15)$$

Using the same weights $\alpha_i(t)$, a time-varying state feedback gain for system (8) is constructed as:

$$u(t) = \tilde{K}_e \int_0^t e(\tau) d\tau + \tilde{\mathbf{K}}_x \mathbf{x}(t) \quad (16)$$

Where:

$$\tilde{K}_e(t) = \sum_{i=1}^N \alpha_i(t) K_{e_i}, \quad \tilde{\mathbf{K}}_x(t) = \sum_{i=1}^N \alpha_i(t) \mathbf{K}_{x_i}$$

Figure 1 illustrates the implementation of the control policy described in equation (15) for each local model, as defined in equation (3). Additionally, Figure 2 depicts the global controller that enables the implementation of the control law in equation (16) by interpolating the local controllers, either using weights computed analytically (equation (5)) or geometrically (equation (7)). For the augmented tracking system in equation (13), the gain $\mathbf{K}_i = \begin{bmatrix} \underbrace{K_{e_i}}_{(1 \times 1)} & \underbrace{\mathbf{K}_{x_i}}_{(1 \times N)} \end{bmatrix}$ is computed for each $i = 1, \dots, N$ by solving a closed-loop pole placement problem using the Matlab command:

$$\gg \mathbf{K}_i = \text{place}(\mathbf{F}_i, \mathbf{G}_i, P) \quad (17)$$

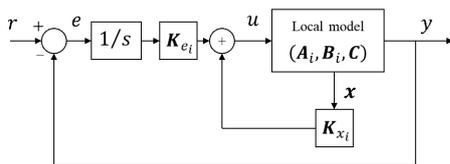


Figure 1. Local tracking control system block diagram

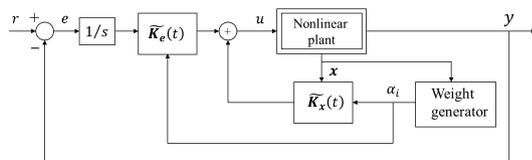


Figure 2. Overall tracking controller implementation by interpolating local controllers

Where P represents the desired closed-loop poles, selected to satisfy a guaranteed pole-dominant criterion [25, 26], based on closed-loop design requirements specified in the time domain, such as overshoot (OS) and settling time (T_s). In light of the above discussion, the design algorithm for implementing the interpolated control law in equation (16) is summarized in Table 1.

Table 1. Design algorithm for quasi-LPV control

Step 1	Construct a quasi-LPV model (2) for the nonlinear system to be controlled (1).
Step 2	From (2), derive a set of local linearized models (3).
Step 3	Compute the local gains k_{e_i} y \mathbf{K}_{x_i} en (15) para cada modelo de seguimiento local aumentado en (13), in (15) for each local tracking augmented model in (13), using the closed-loop specifications for OS y T_s through the Matlab command (17).
Step 4	Compute the weights α_i by continuously solving (5) or (7).
Step 5	Interpolate the local controllers obtained in Step (2) through (16).

2.4. Coupled Tank System

Figure 3 depicts the coupled tank system. It consists of a single pump and two tanks, each equipped with a pressure sensor to measure the water level. The pump transfers water from the bottom reservoir to the top of the system. Depending on the configuration of the outflow valves, water can flow into the upper tank, the lower tank, or both. This configuration is illustrated in Figure 4, where the pump output is connected to the first tank.

x_1 and x_2 represent the water levels in tanks 1 and 2, respectively. The vector functions in the form of equation (18) for the coupled tank system are derived using Bernoulli's law and the mass balance principle [34] and are expressed as:

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} -(Ad_1/A_1 \sqrt{2gx_1(t)}) & 0 \\ (Ad_1/A_2 \sqrt{2gx_1(t)}) & -(Ad_2/A_2 \sqrt{2gx_2(t)}) \end{bmatrix} \quad \mathbf{g}(\mathbf{x}) = \begin{bmatrix} K_f/A_1 \\ 0 \end{bmatrix} \quad (18)$$



Figure 3. Coupled tank system

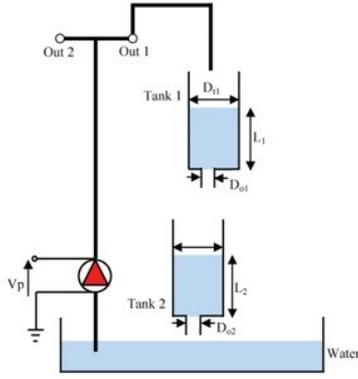


Figure 4. Standard configuration of the coupled tank system

Where A_1 and A_2 denote the cross-sectional areas of tanks 1 and 2, respectively. A_{d1}, A_{d2} represent the cross-sectional areas of the corresponding orifices, g is the acceleration on Earth due to gravity, and K_f is the pump flow constant. The numerical values of these parameters are provided in Table 2.

Table 2. Physical parameters of the coupled tank system

Description	Value	Unit
Pump flow constant (K_f)	4	$cm^3/s/V$
Small Outflow Orifice Diameter of Tank 1 (D_{o1})	0.635	cm
Small Outflow Orifice Diameter of Tank 2 (D_{o2})	0.476	cm
Tanks' Diameter (D_{t1}, D_{t2})	4.445	cm
Water levels range of Tanks 1 and 2	30	cm
Acceleration due to gravity (g)	981	cm/s^2
Pump peak voltage	22	V

3. Results and discussion

This section outlines the implementation and evaluates the performance of the quasi-LPV control method, as summarized in Table 1.

3.1. Quasi-LPV model

The input voltage applied to the pump serves as the control signal, while the water level in the second tank is selected as the controlled output. Based on equation (18), the nonlinear model of the tank system can be expressed as:

$$\dot{x}_1(t) = -\frac{Ad_1}{A_1} \sqrt{2gx_1(t)} + \frac{K_f}{A_1} u(t)$$

$$\dot{x}_2(t) = \frac{Ad_1}{A_2} \sqrt{2gx_1(t)} - \frac{Ad_2}{A_2} \sqrt{2gx_2(t)}$$

The nonlinear terms in each equation can be reformulated as follows:

$$\dot{x}_1(t) = -\frac{Ad_1}{A_1} \sqrt{\frac{2gx_1^2(t)}{x_1(t)}} + \frac{K_f}{A_1} u(t)$$

$$\dot{x}_2(t) = \frac{Ad_1}{A_2} \sqrt{\frac{2gx_1^2(t)}{x_1(t)}} - \frac{Ad_2}{A_2} \sqrt{\frac{2gx_2^2(t)}{x_2(t)}}$$

Resulting in:

$$\dot{x}_1(t) = -\frac{Ad_1\sqrt{2g}}{A_1} \sqrt{\frac{1}{x_1(t)}} x_1(t) + \frac{K_f}{A_1} u(t) \quad (19)$$

$$\dot{x}_2(t) = \frac{Ad_1\sqrt{2g}}{A_2} \sqrt{\frac{1}{x_1(t)}} x_1(t) - \frac{Ad_2\sqrt{2g}}{A_2} \sqrt{\frac{1}{x_2(t)}} x_2(t)$$

Defining the parameter vector in (19) as:

$$\theta(t) = [\theta_1(t) \ \theta_2(t)]^T = [1/\sqrt{x_1} \ 1/\sqrt{x_2}]^T \quad (20)$$

Utilizing the numerical values from Table 2, the quasi-LPV model in the form of equation (2) is expressed as:

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} -0.904\theta_1(t) & 0 \\ 0.904\theta_1(t) & -0.508\theta_2(t) \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0.258 \\ 0 \end{bmatrix} u(t) \quad (21)$$

$$y(t) = \underbrace{[0 \ 1]}_C \mathbf{x}(t)$$

The liquid levels in the tanks are considered uncertain but vary within their physical limits, as specified in Table 2, over the interval:

$$x_1(t), x_2(t) \in [5 \ 25] cm \quad (22)$$

When the liquid levels in the tanks vary within the range specified in equation (22), the parameter vector in equation (20) will fluctuate within the rectangular bounding box:

$$\theta_1(t), \theta_2(t) \in [0.20 \ 0.45] \quad (23)$$

3.2. Local linearized models

The extreme parameter combinations within the bounding box in equation (23) yield the following vectors:

$$\begin{aligned} \theta^1 &= \begin{bmatrix} 0.20 \\ 0.20 \end{bmatrix}, & \theta^2 &= \begin{bmatrix} 0.20 \\ 0.45 \end{bmatrix}, \\ \theta^3 &= \begin{bmatrix} 0.45 \\ 0.20 \end{bmatrix}, & \theta^4 &= \begin{bmatrix} 0.45 \\ 0.45 \end{bmatrix} \end{aligned} \quad (24)$$

This results in the following set of local linearized models: $(\mathbf{A}_i, \mathbf{B}_i) = (\mathbf{A}(\theta^i), \mathbf{B})$ for $i = 1, \dots, 4$.

$$\begin{aligned} \mathbf{A}_1 &= \begin{bmatrix} -0.181 & 0 \\ 0.181 & -0.102 \end{bmatrix} & \mathbf{A}_2 &= \begin{bmatrix} -0.181 & 0 \\ 0.181 & -0.229 \end{bmatrix} \\ \mathbf{A}_3 &= \begin{bmatrix} -0.407 & 0 \\ 0.407 & -0.102 \end{bmatrix} & \mathbf{A}_4 &= \begin{bmatrix} -0.407 & 0 \\ 0.407 & -0.229 \end{bmatrix} \\ \mathbf{B} &= \begin{bmatrix} 0.258 \\ 0 \end{bmatrix} \end{aligned} \quad (25)$$

3.3. Local controllers

Using equation (25), the augmented systems in equation (13) for each vertex are given by:

$$\dot{z}(t) = \begin{bmatrix} 0 & -C \\ \mathbf{0} & \mathbf{A}_i \end{bmatrix} z(t) + \begin{bmatrix} 0 \\ \mathbf{B} \end{bmatrix} u_0(t) \quad (26)$$

The four controller gains \mathbf{K}_i in equation (15) are computed using the closed-loop time domain specifications $OS = 1\%$ and $T_s = 40s$. The calculation is then performed as outlined in [35].

$$\begin{aligned} OS &= e^{(-\zeta\pi/\sqrt{1-\zeta^2})} \Rightarrow \zeta = \frac{1}{\sqrt{1 + \left(\frac{\pi}{\ln(OS)}\right)^2}} \\ &= 0.83 \\ T_s &= 4/\zeta\omega_n \Rightarrow \omega_n = \frac{4}{\zeta T_s} = 0.1 \end{aligned}$$

The resulting dominant poles $p_{1,2} = -0.0996 \pm j0.0669(s^2 + 0.1992s + 0.0144)$. The desired closed-loop poles used in equation (17) are $P = [-0.0996 \pm j0.0669, -0.996]$, where $p_3 = -0.996$ is a fast pole with negligible influence on the OS and T_s specifications. The controller gains are computed using equation (17) as follows:

$$\begin{aligned} \mathbf{K}_1 &= \left[\begin{array}{c|cc} \underbrace{-0.3161}_{K_{e1}} & \underbrace{3.5380 \quad 2.1888}_{K_{x1}} \\ \hline \underbrace{-0.3161}_{K_{e2}} & \underbrace{3.0457 \quad -0.1663}_{K_{x2}} \\ \underbrace{-0.1405}_{K_{e3}} & \underbrace{2.6620 \quad 0.9728}_{K_{x3}} \\ \underbrace{-0.1405}_{K_{e4}} & \underbrace{2.1698 \quad -0.0739}_{K_{x4}} \end{array} \right] \end{aligned} \quad (27)$$

3.4. Interpolation mechanisms

In the analytical approach, equation (5) is represented as the following system of linear equations:

$$\underbrace{\begin{bmatrix} 0.20 & 0.20 & 0.45 & 0.45 \\ 0.20 & 0.45 & 0.20 & 0.45 \\ 1 & 1 & 1 & 1 \end{bmatrix}}_W \begin{bmatrix} \alpha_1(t) \\ \alpha_2(t) \\ \alpha_3(t) \\ \alpha_4(t) \end{bmatrix} = \begin{bmatrix} \theta_1(t) \\ \theta_2(t) \\ 1 \end{bmatrix}$$

Solving this system using the pseudoinverse matrix $(\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T$ yields the following equation:

$$\begin{bmatrix} \alpha_1(t) \\ \alpha_2(t) \\ \alpha_3(t) \\ \alpha_4(t) \end{bmatrix} = \begin{bmatrix} -0.50 & -0.50 & 1 \\ -3.25 & 0.75 & 1 \\ 0.75 & -3.25 & 1 \\ 3.50 & 3.50 & -2 \end{bmatrix} \begin{bmatrix} \theta_1(t) \\ \theta_2(t) \\ 1 \end{bmatrix} \quad (28)$$

For geometric interpolation, equation (7) is implemented directly using a Matlab function block. A straightforward Matlab function code is written and integrated into a Simulink model, which executes the simulation.

3.5. Gain-scheduled control implementation

The gain-scheduled control strategy depicted in Figure 2 was implemented. Figure 5 illustrates the liquid level response in the second tank following a set-point change, comparing both interpolation methods for the computed linear controllers (27).

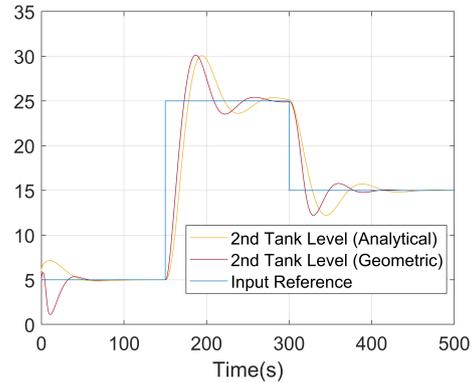


Figure 5. Second tank closed-loop liquid level response for analytical and geometric interpolation methods

The geometric method encounters specific issues at the start of the simulation due to its inability to provide the required negative control action. After this initial phase, the performance of both interpolation schemes becomes comparable.

Figure 6 illustrates the control signal, while Figure 7 focuses on the first 20 seconds of the control signal.

It is evident that when a negative control signal is required, the geometric scheme remains at zero, confirming the issues observed at the beginning of the simulation, as depicted in Figure 5. It is important to note that the control signal provided by the pump cannot be negative, a limitation not accounted for during the simulation when evaluating the performance of both interpolation schemes.

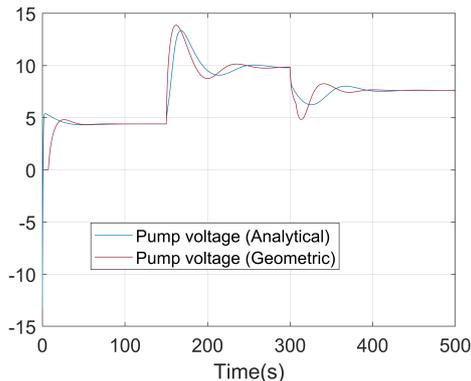


Figure 6. Pump voltage control signal for analytical and geometric interpolation methods

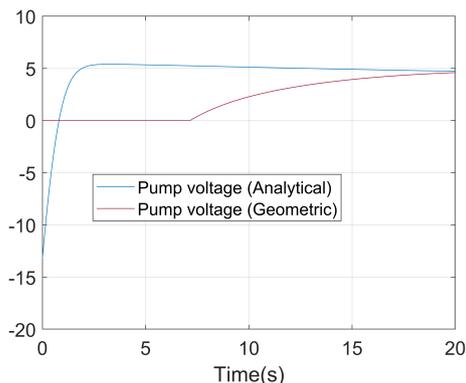


Figure 7. Detail of the control signal during the first 20 s

3.6. Further results

Figure 8 illustrates the parametric bounding box defined by equation (23). The previous results involved the implementation of the gain scheduled controller through the interpolation both geometric and analytical of the local controllers computed at the vertices (A), (B), (C), and (D), based on a dominant pole criterion for the desired OS and T_s specifications. Additionally, the simulation permitted the control signal to take on negative values to facilitate a comparison between the two interpolation mechanisms.

The gain-scheduled controller is implemented in this section, using various local controllers computed within the region shown in Figure 8, as specified in

Table 3. The control signal is constrained to remain within the operational range of the pump (0-22 V), and the desired closed-loop poles in equation (17) are selected as $P = [-0.1, -0.2, -10]$, rather than employing the dominant pole criterion.

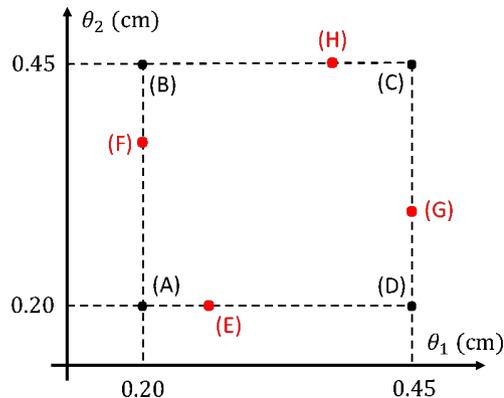


Figure 8. Points chosen in the parametric bounding box (23) to compute local controllers

Table 3. Points chosen in region (3) to compute local controllers.

Model	Points chosen in (23)
M1	(A), (B), (C), (D)
M2	(E), (F), (G), (H)
M3	(A), (D)

Model M1 utilizes the vertices of the region, M2 computes the local controllers along the edges, and model M3 considers the extreme vertices of the region, where parameters θ_1 and θ_2 take their minimum and maximum possible values. The selection of the M3 model is justified by the well-known Edge Theorem [36]. Figure 9 illustrates the level in the second tank and the pump control signal using the M1 model with analytical interpolation.

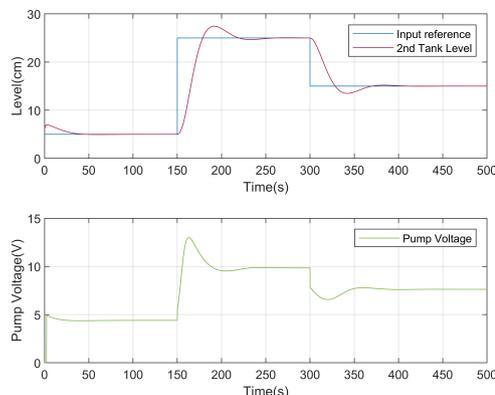


Figure 9. Second tank level and pump voltage for M1 model and analytical interpolation

Figure 10 presents a similar scenario employing geometric interpolation. Figures 11 , 12 replicate the analysis for the M2 model, using analytical and geometric interpolation mechanisms, respectively.

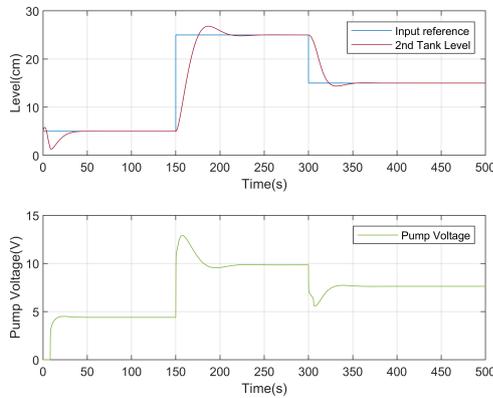


Figure 10. Second tank level and pump voltage for M1 model and geometric interpolation

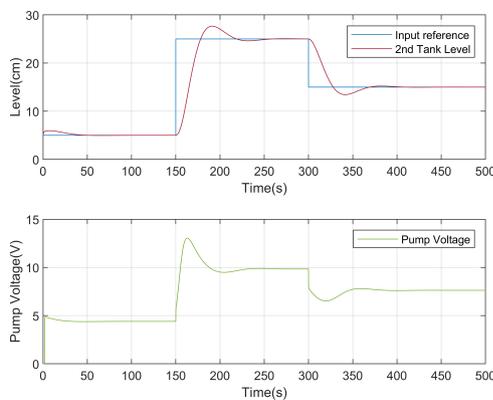


Figure 11. Second tank level and pump voltage for M2 model and analytical interpolation

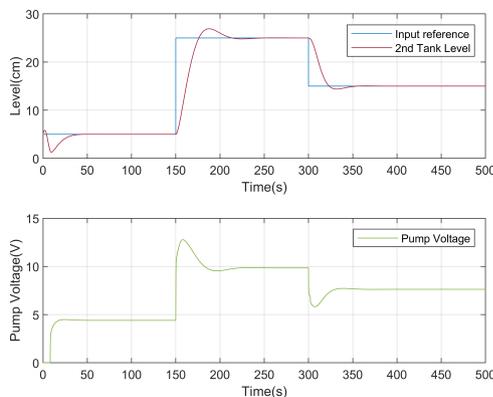


Figure 12. Second tank level and pump voltage for M2 model and geometric interpolation

Figures 13 and 14 display the results for the M3 model, again using analytical and geometric interpolation, respectively. Finally, Figures 15 and 16 compare the evolution of the liquid level in the second tank for all three models, with analytical and geometric interpolation considered, respectively.

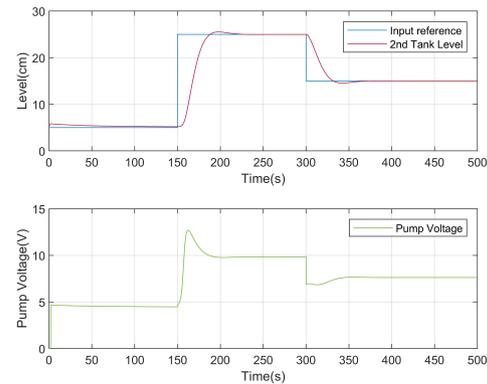


Figure 13. Second tank level and pump voltage for M3 model and analytical interpolation

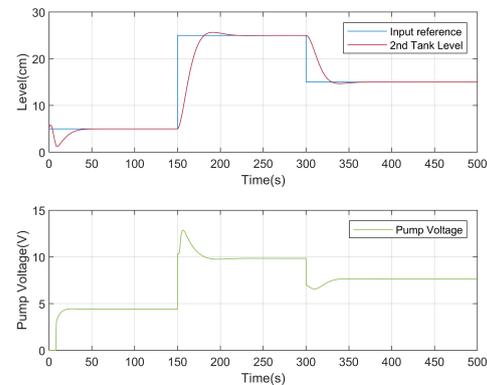


Figure 14. Second tank level and pump voltage for M3 model and geometric interpolation

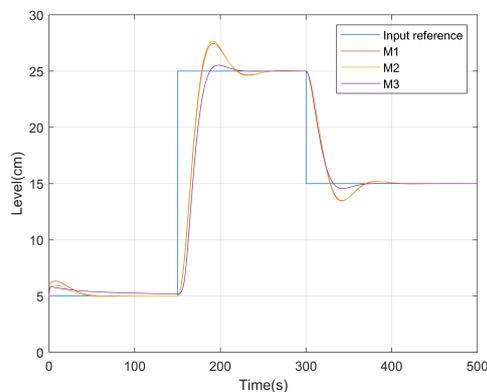


Figure 15. Second tank level for models M1, M2 and M3 using analytical interpolation

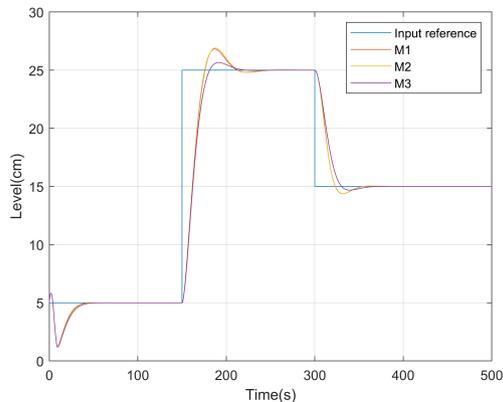


Figure 16. Second tank level for models M1, M2 and M3 using geometric interpolation

4. Conclusions

A gain-scheduled procedure was proposed to control a coupled tank system modeled as a quasi-LPV system. The nonlinearities of the model are directly captured by a set of uncertain parameters that vary within a bounded set, constrained by the physical limits of the tank system. Extreme combinations of the parameter vector were computed, and local linear approximations were obtained. These approximations were then used in the state-space synthesis of control laws to track a constant reference input. The global controller was constructed as a weighted average of the local contributions, where the weights depended on the instantaneous values of the parameter vector. Two interpolation mechanisms, geometric and analytical, were employed to determine the weighted average of the linear controllers. The geometric method is based on the Euclidean distance between the parameter vector and the vertices, while the analytical method involves solving a linear system of equations using the pseudoinverse of a matrix. The geometric scheme is simpler and generates only positive control actions, with a very short computation time. In contrast, the analytical scheme can provide both positive and negative control actions but requires significantly more processing time. Simulation results demonstrated that using the two extreme vertices (Model M3) to compute the interpolated local controllers reduces the computational effort needed.

The primary limitation of the methodology is the challenge of accurately determining the quasi-LPV model to capture the system's nonlinearities, which is not an easy task for all plants. This indicates that the proposed approach may not be universally applicable. However, when a nonlinear plant can be effectively modeled using a linear parameter-varying system, the method is straightforward to implement and yields satisfactory results. Another key aspect

of the method is that the control law for designing the local controllers is not limited to closed-loop pole assignment, as demonstrated in this article. Various state-feedback control strategies can be employed, including those that account for optimality, robustness, and constraints. Additionally, although the guaranteed pole-dominant criterion is suitable for linear systems, a notable discrepancy emerged between the design specifications and the actual performance in the case of the nonlinear tank system. This gap was mitigated by setting dominant real poles to improve control over the output.

Ongoing work focuses on the real-time implementation of the proposed design method and the inclusion of state observers.

Acknowledgements

The authors gratefully acknowledge the support provided by the Research Program of the Metropolitan University in Caracas, Venezuela, under project number PG-A-13-21-22.

References

- [1] R. Grygiel, R. Bieda, and M. Blachuta, "On significance of second-order dynamics for coupled tanks systems," in *2016 21st International Conference on Methods and Models in Automation and Robotics (MMAR)*, 2016, pp. 1016–1021. [Online]. Available: <https://doi.org/10.1109/MMAR.2016.7575277>
- [2] S. K. Singh, N. Katal, and S. G. Modani, "Multi-objective optimization of pid controller for coupled-tank liquid-level control system using genetic algorithm," in *Proceedings of the Second International Conference on Soft Computing for Problem Solving (SocProS 2012), December 28-30, 2012*, B. V. Babu, A. Nagar, K. Deep, M. Pant, J. C. Bansal, K. Ray, and U. Gupta, Eds. New Delhi: Springer India, 2014, pp. 59–66. [Online]. Available: https://doi.org/10.1007/978-81-322-1602-5_7
- [3] D. L. Mute, S. R. Mahapatro, and K. K. Chaudhari, "Internal model based pi controller design for the coupled tank system: An experimental study," in *2016 IEEE First International Conference on Control, Measurement and Instrumentation (CMI)*, 2016, pp. 72–76. [Online]. Available: <https://doi.org/10.1109/CMI.2016.7413713>
- [4] M. G. Stohy, H. S. Abbas, A.-H. M. El-Sayed, and A. G. Abo El-maged, "Parameter estimation and pi control for a water coupled tank system," *Journal of Advanced Engineering Trends*, vol. 38,

- no. 2, pp. 147–159, 2020. [Online]. Available: <https://doi.org/10.21608/jaet.2020.73062>
- [5] L. Liang, “The application of fuzzy pid controller in coupled-tank liquid-level control system,” in *2011 International Conference on Electronics, Communications and Control (ICECC)*, 2011, pp. 2894–2897. [Online]. Available: <https://doi.org/10.1109/ICECC.2011.6067785>
- [6] A. Başçı and A. Derdiyok, “Implementation of an adaptive fuzzy compensator for coupled tank liquid level control system,” *Measurement*, vol. 91, pp. 12–18, 2016. [Online]. Available: <https://doi.org/10.1016/j.measurement.2016.05.026>
- [7] M. Essahafi, “Model predictive control (MPC) applied to coupled tank liquid level system,” *CoRR*, vol. abs/1404.1498, 2014. [Online]. Available: <https://doi.org/10.48550/arXiv.1404.1498>
- [8] J. Berberich, J. Köhler, M. A. Müller, and F. Allgöwer, “Data-driven model predictive control: closed-loop guarantees and experimental results,” *at - Automatisierungstechnik*, vol. 69, no. 7, pp. 608–618, Jun. 2021. [Online]. Available: <http://dx.doi.org/10.1515/auto-2021-0024>
- [9] J. Jiffy Anna, N. E. Jaffar, and F. Riya Mary, “Modelling and control of coupled tank liquid level system using backstepping method,” *International Journal of Engineering Research and Technology*, vol. 4, no. 6, pp. 667–671, 2015. [Online]. Available: <http://dx.doi.org/10.17577/IJERTV4IS060710>
- [10] J. Dai, B. Ren, and Q.-C. Zhong, “Uncertainty and disturbance estimator-based backstepping control for nonlinear systems with mismatched uncertainties and disturbances,” *Journal of Dynamic Systems, Measurement, and Control*, vol. 140, no. 12, Jul 2018, 121005. [Online]. Available: <https://doi.org/10.1115/1.4040590>
- [11] F. Abu Khadra and J. Abu Qudeiri, “Second order sliding mode control of the coupled tanks system,” *Mathematical Problems in Engineering*, vol. 2015, no. 1, p. 167852, 2015. [Online]. Available: <https://doi.org/10.1155/2015/167852>
- [12] K. K. Ayten and A. Dumlu, “Implementation of a pid type sliding-mode controller design based on fractional order calculus for industrial process system,” *Elektronika ir Elektrotechnika*, vol. 27, no. 6, pp. 4–10, Dec. 2021. [Online]. Available: <https://doi.org/10.5755/j02.eie.30306>
- [13] M. Vashishth, L. S. Rai, and A. V. R. Kumar, “Liquid level control of coupled tank system using fractional pid controller,” 2013. [Online]. Available: <https://upsalesiana.ec/ing33ar2r13>
- [14] P. Roy and B. Krishna Roy, “Fractional order pi control applied to level control in coupled two tank mimo system with experimental validation,” *Control Engineering Practice*, vol. 48, pp. 119–135, 2016. [Online]. Available: <https://doi.org/10.1016/j.conengprac.2016.01.002>
- [15] M. H. Jali, A. Ibrahim, R. Ghazali, C. C. Soon, and A. R. Muhammad, “Robust Control Approach of SISO Coupled Tank System,” *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 1, 2021. [Online]. Available: <http://dx.doi.org/10.14569/IJACSA.2021.0120123>
- [16] P. Teppa, M. Bravo, and G. Garcia, “Control por rechazo activo de perturbaciones del nivel de líquido de un sistema de tanques acoplado,” *Revista Faraute de Ciencia y Tecnología*, vol. 7, no. 1, pp. 10–18, 2012. [Online]. Available: <https://upsalesiana.ec/ing33ar2r16>
- [17] P. Teppa Garran and G. Garcia, “Design of an optimal pid controller for a coupled tanks system employing adrc,” *IEEE Latin America Transactions*, vol. 15, no. 2, pp. 189–196, 2017. [Online]. Available: <https://doi.org/10.1109/TLA.2017.7854611>
- [18] P. Teppa Garran, M. Faggioni, and G. Garcia, “Optimal tracking in two-degree-of-freedom control systems: Coupled tank system,” *Journal of Applied Research and Technology*, vol. 21, no. 4, pp. 560–570, 2023. [Online]. Available: <https://doi.org/10.22201/icat.24486736e.2023.21.4.2000>
- [19] W. J. Rugh and J. S. Shamma, “Research on gain scheduling,” *Automatica*, vol. 36, no. 10, pp. 1401–1425, 2000. [Online]. Available: [https://doi.org/10.1016/S0005-1098\(00\)00058-3](https://doi.org/10.1016/S0005-1098(00)00058-3)
- [20] D. Rotondo, *Advances in Gain-Scheduling and Fault Tolerant Control Techniques*. Springer Cham, 2018. [Online]. Available: <https://doi.org/10.1007/978-3-319-62902-5>
- [21] G. S. X. Bombois, D. Ghosh and J. Huillery, “Lpv system identification for control using the local approach,” *International Journal of Control*, vol. 94, no. 2, pp. 390–410, 2021. [Online]. Available: <https://doi.org/10.1080/00207179.2019.1596316>
- [22] Z. GAO and J. FU, “Robust lpv modeling and control of aircraft flying through wind disturbance,” *Chinese Journal of Aeronautics*, vol. 32, no. 7, pp. 1588–1602, 2019. [Online]. Available: <https://doi.org/10.1016/j.cja.2019.03.029>

- [23] S. Li, A.-T. Nguyen, T.-M. Guerra, and A. Kruszewski, “Reduced-order model based dynamic tracking for soft manipulators: Data-driven lpv modeling, control design and experimental results,” *Control Engineering Practice*, vol. 138, p. 105618, 2023. [Online]. Available: <https://doi.org/10.1016/j.conengprac.2023.105618>
- [24] R. Robles, A. Sala, and M. Bernal, “Performance-oriented quasi-lpv modeling of nonlinear systems,” *International Journal of Robust and Nonlinear Control*, vol. 29, no. 5, pp. 1230–1248, 2019. [Online]. Available: <https://doi.org/10.1002/rnc.4444>
- [25] P. Persson and K. Astrom, “Dominant pole design - a unified view of pid controller tuning,” *IFAC Proceedings Volumes*, vol. 25, no. 14, pp. 377–382, 1992. [Online]. Available: [https://doi.org/10.1016/S1474-6670\(17\)50763-6](https://doi.org/10.1016/S1474-6670(17)50763-6)
- [26] A. D. Mammadov, E. Dincel, and M. T. Söylemez, “Analytical design of discrete pi-pr controllers via dominant pole assignment,” *ISA Transactions*, vol. 123, pp. 312–322, 2022. [Online]. Available: <https://doi.org/10.1016/j.isatra.2021.05.038>
- [27] P. Teppa-Garrán and J. Andrade-Da Silva, “An analytical interpolation approach for gain scheduling control of linear parameter varying systems,” *Ciencia e Ingeniería*, vol. 30, no. 1, pp. 93–101, 2008. [Online]. Available: <https://upsalesiana.ec/ing33ar2r27>
- [28] W. Tan, A. Packard, and G. Balas, “Quasi-lpv modeling and lpv control of a generic missile,” in *Proceedings of the 2000 American Control Conference. ACC (IEEE Cat. No.00CH36334)*, vol. 5, 2000, pp. 3692–3696. [Online]. Available: <https://doi.org/10.1109/ACC.2000.879259>
- [29] G. Vinco, S. Theodoulis, O. Sename, and G. Strub, “Quasi-lpv modeling of guided projectile pitch dynamics through state transformation technique,” *IFAC-PapersOnLine*, vol. 55, no. 35, pp. 43–48, 2022, 5th IFAC Workshop on Linear Parameter Varying Systems LPVS 2022. [Online]. Available: <https://doi.org/10.1016/j.ifacol.2022.11.288>
- [30] Z. He and L. Zhao, “Quadrotor trajectory tracking based on quasi-lpv system and internal model control,” *Mathematical Problems in Engineering*, vol. 2015, no. 1, p. 857291, 2015. [Online]. Available: <https://doi.org/10.1155/2015/857291>
- [31] P. S. Cisneros, A. Sridharan, and H. Werner, “Constrained predictive control of a robotic manipulator using quasi-lpv representations,” *IFAC-PapersOnLine*, vol. 51, no. 26, pp. 118–123, 2018, 2nd IFAC Workshop on Linear Parameter Varying Systems LPVS 2018. [Online]. Available: <https://doi.org/10.1016/j.ifacol.2018.11.158>
- [32] H. Xie, L. Dai, Y. Lu, and Y. Xia, “Disturbance rejection mpc framework for input-affine nonlinear systems,” *IEEE Transactions on Automatic Control*, vol. 67, no. 12, pp. 6595–6610, 2022. [Online]. Available: <https://doi.org/10.1109/TAC.2021.3133376>
- [33] J. Stefanovski, “Feedback affinization of nonlinear control systems,” *Systems & Control Letters*, vol. 46, no. 3, pp. 207–217, 2002. [Online]. Available: [https://doi.org/10.1016/S0167-6911\(02\)00136-6](https://doi.org/10.1016/S0167-6911(02)00136-6)
- [34] J. Apkarian, *Coupled water tank experiments manual*. Quanser Consulting Advanced Teaching Systems, 1999. [Online]. Available: <https://upsalesiana.ec/ing33ar2r34>
- [35] R. C. Dorf and R. H. Bishop, *Modern Control Systems*. Prentice Hall, 2008. [Online]. Available: <https://upsalesiana.ec/ing33ar2r35>
- [36] A. Bartlett, C. Hollot, and H. Lin, “Root locations of an entire polytope of polynomials: It suffices to check the edges,” in *1987 American Control Conference*, 1987, pp. 1611–1616. [Online]. Available: <https://doi.org/10.23919/ACC.1987.4789571>



APPLICATION OF MANETs AS A COMMUNICATION SYSTEM FOR SUSTAINABLE MOBILITY

APLICACIÓN DE MANET COMO SISTEMA DE COMUNICACIÓN EN LA MOVILIDAD SOSTENIBLE

Nancy Eras¹ , José Andrés Otavalo¹ , Santiago González^{1,*}

Received: 09-05-2024, Received after review: 10-06-2024, Accepted: 16-09-2024, Published: 01-01-2025

Abstract

This paper presents an architecture based on the MANET (Mobile Ad Hoc Network) paradigm as an emergency communication system between users of electric bicycles. The solution consists of 4 mobile nodes representing the users and a main fixed node, which emulates a bicycle docking station. This architecture allows multi-hop communication between the nodes, using the proactive routing protocols OLSR (Optimized Link State Routing) and BATMAN (Better Approach to Mobile Ad Hoc Networking). The study was divided into 3 main stages. First, an analysis of the wireless medium was performed to determine the maximum transmission distance and the maximum bitrate between 2 nodes. Subsequently, the throughput behavior was characterized in a multi-hop configuration consisting of 4 nodes in order to establish the network capacity in terms of bandwidth. Finally, a web application was implemented for the transmission of audio and text traffic. Regarding the evaluation of the proposal, two scenarios were designed to emulate the integration of a new cyclist to the network and the communication between two users in motion. The results reveal that OLSR provides a better system operation, with a throughput of 2.54 Mbps at 3 hops and a PRR (Packet Reception Rate) higher than 96%. In addition, it guarantees a delay within the ITU-T (International Telecommunication Union-Telecommunication) G.114 recommendation for bidirectional communication.

Keywords: BATMAN, BSS, Emergency communication system, ITS, MANET, OLSR

Resumen

En este artículo se presenta una arquitectura basada en el paradigma MANET (Mobile Ad Hoc Network) como un sistema de comunicación de emergencia entre usuarios de bicicletas eléctricas. La solución consta de cuatro nodos móviles que representan a los usuarios y un nodo fijo principal, que emula una estación de anclaje de bicicletas. Esta arquitectura permite la comunicación multisalto entre los nodos, utilizando los protocolos de enrutamiento proactivos OLSR (Optimized Link State Routing) y BATMAN (Better Approach to Mobile Ad Hoc Networking). El estudio se dividió en tres etapas principales. Primero, se hizo un análisis del medio inalámbrico para determinar la distancia máxima de transmisión y el bitrate máximo entre dos nodos. Posteriormente, se caracterizó el comportamiento del throughput en una configuración multisalto conformada por cuatro nodos con el fin de establecer la capacidad de la red en términos de ancho de banda. Finalmente, se implementó una aplicación web para la transmisión de tráfico de audio y texto. En cuanto a la evaluación de la propuesta, se diseñaron dos escenarios que emulan la integración de un nuevo ciclista a la red y la comunicación entre dos usuarios en movimiento. Los resultados revelan que OLSR proporciona una mejor operación del sistema, con un throughput de 2.54 Mbps a 3 saltos y un PRR (*Packet Reception Rate*) superior al 96 %. Además, garantiza un delay dentro de la recomendación G.114 de la ITU-T (*International Telecommunication Union-Telecommunication*) para una comunicación bidireccional.

Palabras clave: BATMAN, BSS, ITS, MANET, OLSR, sistema de comunicación de emergencia

^{1,*}Departamento de Ingeniería Eléctrica, Electrónica y Telecomunicaciones, Universidad de Cuenca, Cuenca, Ecuador.
 Corresponding Author✉: santiago.gonzalezm@ucuenca.edu.ec.

Suggested citation: Eras, N.; Otavalo, J. A. and González, S. "Application of MANETs as a communication system for sustainable mobility," *Ingenius, Revista de Ciencia y Tecnología*, N.º 33, pp. 27-37, 2025, DOI: <https://doi.org/10.17163/ings.n33.2025.03>.

1. Introduction

The bicycle is a vital mode of transportation for the development of sustainable mobility systems. As highlighted in [1], it offers users numerous advantages, including affordability, efficiency, safety, and environmental sustainability. Consequently, many cities have implemented Bike Sharing Systems (BSS) to promote cycling as a viable transportation option. Nevertheless, despite these benefits, users remain exposed to challenges such as traffic congestion, accidents, environmental pollution, and noise pollution, among other adverse conditions [2].

In this context, Bike Sharing Systems (BSS) have undergone significant technological advancements and are now in their fourth and fifth generations. These iterations include enhancements to streamline sharing processes, integrate electric bicycles, and implement communication systems among users [3]. As part of Intelligent Transport Systems (ITS), BSS have been widely adopted in cities worldwide, contributing significantly to reducing CO_2 emissions [4–8].

On the other hand, considering that a BSS involves multiple users, ITS systems have incorporated hybrid communication architectures based on the Ad Hoc paradigm, particularly leveraging ANET (Mobile Ad Hoc Network) and VANET (Vehicular Ad Hoc Network).

In MANET-type networks, each node functions both as a terminal device and as a router, enabling the rapid establishment of communication links without relying on centralized network infrastructure. As a result, they present a valuable technological solution for scenarios requiring resilient communication systems, as highlighted in [9, 10]. However, their implementation poses several challenges, including the management of dynamic topologies due to node mobility, energy limitations in battery-powered devices, and the variability of the wireless medium, particularly in multi-hop configurations [11, 12].

In this context, the literature has proposed various routing mechanisms to address the challenges associated with diverse applications and scenarios, including MANETs, VANETs, FANETs (Flying Ad Hoc Networks), and SANETs (Sea Ad Hoc Networks), among others [13]. Notably, prior studies emphasize the superior functionality of proactive protocols compared to reactive or hybrid approaches, even in environments characterized by highly dynamic topologies [14, 15].

It is important to note that most proposals have been evaluated within simulation environments and under controlled conditions, highlighting the additional challenges and complexities associated with experimentation in real-world scenarios and applications [14]. The most relevant studies available in the literature are discussed below, with a particular focus on applications requiring resilient or emergency communication

systems.

Emergency communication systems play a vital role in scenarios where conventional telecommunication infrastructures are non-operational, such as during earthquakes, floods, accidents, or in remote and low-coverage areas [12, 16].

In [17], an evaluation of routing protocols in emergency applications using FANETs is presented. The results highlight the performance of AODV (Ad Hoc On-Demand Distance Vector), DSDV (Destination Sequenced Distance Vector), and OLSR (Optimized Link State Routing) mechanisms. Conversely, [18] emphasizes that proactive protocols in emergency applications exhibit lower end-to-end delays, as their routing tables are continuously updated to reflect changes in network topology. However, this advantage introduces additional challenges, including increased bandwidth and energy consumption requirements [19].

In [20], the performance of the OLSR and BATMAN protocols is evaluated across the 2.4 GHz and 5 GHz frequency bands. The results show that OLSR outperforms BATMAN in terms of throughput, achieving 0.91 Mbps at 2.4 GHz and 0.82 Mbps at 5 GHz over a distance of 50 meters. Additionally, OLSR exhibits superior performance in packet loss ratio (PLR), with 11% for 2.4 GHz and 20.4% for 5 GHz.

Similarly, in [21], a comparative analysis of the OLSR and BATMAN protocols is presented, defining two distinct scenarios. The first involves a multi-hop topology with four static nodes, where OLSR demonstrates higher throughput (10 Mbps at two hops and 6 Mbps at three hops), while BATMAN exhibits lower delay (6 ms at two hops and 8 ms at three hops). The second scenario involves an evaluation using a mesh topology, where both protocols achieve a 100% packet reception rate.

Multimedia content currently dominates internet traffic. In this context, a comparative analysis of various video codecs over a multi-hop Ad Hoc network is presented in [10]. Key performance metrics such as Packet Reception Rate (PRR), delay, and throughput are evaluated for real-time audio and video transmission. The study concludes that the VP8 video codec is the most suitable for the proposed scenario.

Other studies propose innovative protocol adaptations to ensure the efficient transmission of multimedia traffic. For instance, [22] describes the RTMC (Real-Time Multi-Cast) protocol, which enables receivers to specify their real-time constraints, thereby optimizing multicast trees to meet delay requirements while reducing energy consumption. This solution operates without relying on network topology information or link maintenance mechanisms, offering a reactive routing approach.

On the other hand, [23] introduces the Multi-Parameter Fuzzy Logic Resource Management (MP-FLRM) approach, which leverages request, download,

and upload time data to dynamically update the resource list in real time. This methodology enhances resource management efficiency in VANET scenarios.

Finally, [24] explores the use of the OLSR protocol as a routing mechanism for VoIP (Voice over IP) services in a VANET, aiming to maintain an acceptable quality of service for voice calls. The results indicate that OLSR achieves delay, jitter, and packet loss values of 102.48 ms, 10.675 ms, and 0.07%, respectively. However, limitations are observed as the number of hops increases, resulting in delays that exceed the thresholds recommended by ITU-T standards [15, 25].

Regarding energy constraints, [26] introduces a mechanism known as EARVRT (Energy-Aware Virtual Relay Tunnels), which proposes the establishment of virtual tunnels for route selection based on the available energy in nodes and the number of hops. Additionally, [27] describes the EEE-SR (Enhanced Energy-Efficient Secure Routing) protocol, which integrates security policies, authentication, and energy thresholds into routing decisions.

In [28], a mechanism called ACEAMR (Adaptive Congestion and Energy-Aware Multipath Routing) is introduced, aiming to balance quality of service and energy consumption by discovering stable routes. Simulation results indicate that ACEAMR outperforms existing schemes in throughput (0.2 Kbps at a speed of 15 m/s and 7.2 Kbps at 60 m/s), packet delivery ratio (PDR) (81% at 15 m/s and 63% at 60 m/s), delay (rising from 0.003 ms to 0.0093 ms), and energy efficiency (7.2 J of energy consumption).

Motivated by these considerations, this study proposes a communication system based on a MANET for transmitting emergency informational messages in text and audio formats among BSS users. The system was implemented and evaluated in a real-world scenario, consisting of four mobile nodes and a main node, which facilitates the execution of experiments across all nodes. For the experimental evaluation, the OLSR and BATMAN protocols, highlighted in previous studies, were selected. A comparative analysis of these protocols was performed, focusing on throughput, delay, and packet reception ratio (PRR). Additionally, energy consumption and node autonomy during audio communication operations were analyzed. The primary contribution of this work lies in the experimental analysis and the development of a communication system for BSS utilizing the MANET paradigm in a real-world setting.

The article is organized into several sections, each addressing distinct aspects of the study. The Materials and Methods section outlines the methodology employed for system evaluation, which comprises three key components: the characterization of the Ad Hoc network, the analysis of throughput in a multi-hop topology, and the development and evaluation of a web application. At each stage, tables and figures are

included to enhance clarity and comprehension. In the Results and Discussion section, evaluation metrics such as throughput, delay, and packet reception ratio (PRR) are analyzed and interpreted across two scenarios. The Conclusions section summarizes the study's primary contributions, discusses its practical implications, and provides recommendations for future research directions. Finally, the References section offers the necessary citations to substantiate the work and its findings.

2. Materials and methods

Figure 1 illustrates the coupling of a node with the bicycle, highlighting its main components. Each node is specifically implemented using a Raspberry Pi platform, a wireless card compliant with the IEEE (Institute of Electrical and Electronics Engineers) 802.11 standard and compatible with Ad Hoc mode, as well as a current sensor.

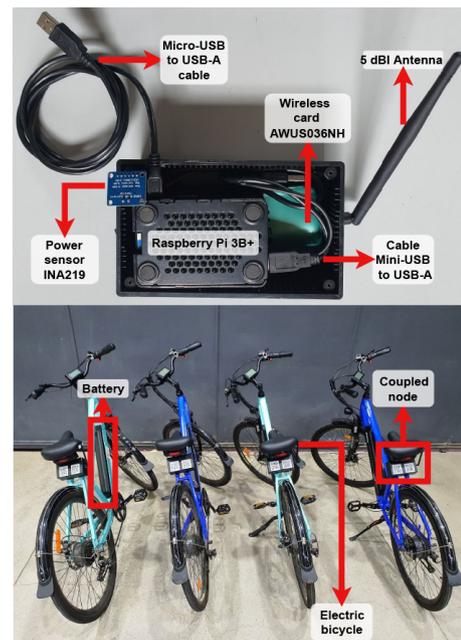


Figure 1. Electric bicycles with the coupled nodes

Figure 2 illustrates the methodology developed to perform the comparative evaluation of routing protocols in the MANET network. Initially, a series of experiments were conducted to characterize the maximum transmission distance and the wireless channel capacity in a two-node configuration. Subsequently, the bandwidth performance was assessed in a linear topology comprising four nodes (three hops). Following this, a mobile web application was developed, enabling both subjective and objective analyses of audio traffic quality. These analyses facilitated the adjustment of audio encoding parameters, tailored to the specific characteristics of the network.

Finally, real-time audio transmission was conducted to verify communication between the nodes, and the system's performance was objectively assessed using metrics such as delay, PRR, and throughput.

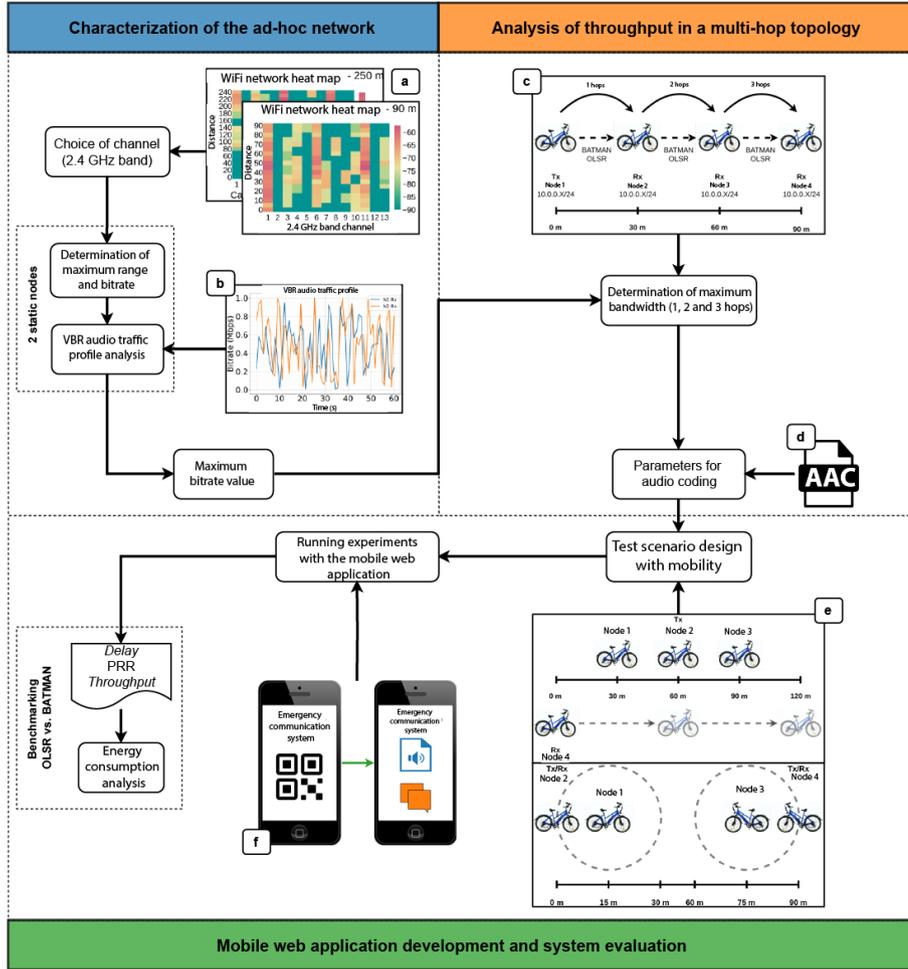


Figure 2. Methodology for evaluating the proposed system: (a) Heat maps of WiFi networks for each channel in the 2.4 GHz band. (b) Graphs representing the Variable Bit Rate VBR audio traffic profile. (c) A multi-hop network topology consisting of four nodes. (d) Advanced Audio Codec (AAC). (e) Test scenarios incorporating mobility to assess network performance and the mobile web application. (f) A mobile web application capable of scanning QR codes and providing audio and text transmission functionalities.

2.1. Characterization of the ad hoc network

The characterization of the Ad Hoc network was conducted to determine the maximum transmission distance and the maximum bitrate for a two-node configuration. The experiments were carried out at the facilities of the Scientific, Technological, and Research Center Balzay (CTI-B) at the University of Cuenca. Specifically, the paths highlighted in Figure 3 were selected, with lengths of 90 meters and 250 meters.

The first experiment involved transmitting UDP (User Datagram Protocol) traffic between two nodes, with one node remaining stationary while the other was moved away in 10-meter increments. At each distance, the traffic rate was maintained at a constant 200 Kbps and repeated ten times. In the second exper-

iment, the nodes were positioned at an intermediate distance, and the UDP traffic rate was incrementally increased from 200 Kbps to 6 Mbps in steps of 100 Kbps. The experiments were conducted using the Iperf tool [29]. Table 1 provides a summary of the main parameters configured for each routing protocol.

Table 1. Default time intervals for the OLSR and BATMAN protocols

OLSR	(s)	BATMAN	(s)
HELLO_INTERVAL	2.0	ORIG_INTERVAL	1.0
REFRESH_INTERVAL	2.0	ELP_INTERVAL	0.5
TC_INTERVAL	5.0		
MID_INTERVAL	TC		
HNA_INTERVAL	TC		



Figure 3. Paths defined for conducting experiments at the CCTI-B.

An additional experiment was conducted to analyze the characteristics and traffic profile generated by audio transmission. For this test, two nodes were positioned 10 meters apart. The AAC (Advanced Audio Coding) codec was employed, configured with an average compression rate of 200 Kbps. The wireless card was set to a transmission speed of 54 Mbps and a power output of 20 dBm. Notably, the FFMPEG (Fast Forward MPEG) tool [30] was utilized for this process.

2.2. Throughput analysis in a multi-hop topology

In this scenario, throughput behavior was analyzed in a multi-hop configuration. This setup is particularly relevant as it enables bicycle users to communicate over extended distances (e.g., between distant nodes or with the anchor station).

The bitrate was varied to assess channel performance across one-hop, two-hop, and three-hop configurations. The experiments involved four nodes arranged in a linear topology, with each node configured to receive traffic exclusively from its directly adjacent node.

Figure 4 illustrates the topology employed in the experiment, where Node 1 serves as the transmitter (TX), and Nodes 2, 3, and 4 function as receivers (RX). The Iperf tool was utilized to regulate UDP traffic in 100 Kbps increments until a threshold value was reached at each hop, following the parameter values outlined in Table 1.

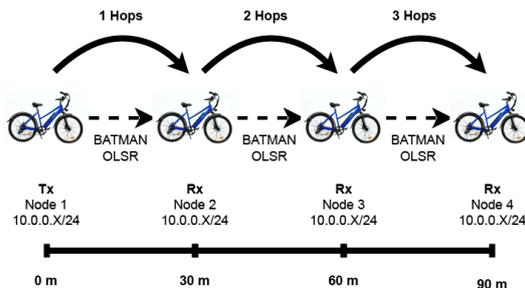


Figure 4. Multi-hop network topology

Based on the results obtained, the audio traffic encoding parameters were adjusted and are presented in detail in Table 2.

Table 2. General parameters for audio transmission and reception

Parameter	Value
Audio codec	AAC
Bitrate	64 kbps
Probesize	50000
Samplerate	48000 Hz

2.3. Development of the mobile web application and system evaluation

At this stage, a mobile web application was developed to support the remote management of the nodes and to evaluate their performance.

The application functions as a DHCP (Dynamic Host Configuration Protocol) server using Hostapd and Dnsmasq, enabling the management of the local network and the assignment of IPv4 (Internet Protocol version 4) addresses to devices connected through a Raspberry Pi.

Additionally, two scenarios were designed to evaluate the MANET network alongside the mobile web application. These scenarios are detailed in the following section.

To enhance user experience, two QR (Quick Response) codes were integrated one to facilitate wireless network connection and the other to launch the mobile web application. These codes enable users to conveniently access the desired functions by scanning them. In cases where scanning capabilities are unavailable, the required information for manually establishing the connection is also provided.

The evaluation was conducted using key metrics, including delay, PRR, throughput, and energy consumption.

3. Results and Discussion

Figure 5 illustrates the bitrate behavior as a function of the number of hops, based on the routing protocol utilized. For the experiments, a linear topology was configured with nodes spaced 30 meters apart. As observed, the throughput decreases as the number of nodes increases, which in turn raises the number of required hops.

This behavior can be attributed to data flow contention at each hop, a phenomenon known as intra flow interference. Additionally, signaling mechanisms, such as acknowledgment (ACK) messages, further limit data transfer speeds, as discussed in [31].

Figure 5a depicts the bitrate performance for the OLSR protocol. For the 1-hop configuration (blue), the

channel exhibits a highly favorable response, demonstrating linear behavior up to 6 Mbps. In the 2-hop scenario (orange), linear behavior persists up to 3 Mbps; however, beyond this point, the channel begins to show variability, with a maximum bitrate of 3.2 Mbps. Lastly, in the 3-hop configuration (green), the behavior remains linear up to 2.4 Mbps. Beyond this threshold, the received bitrate exhibits notable variability, fluctuating between 2.3 Mbps and 2.8 Mbps. The experiment achieved a maximum traffic transmission of 3.3 Mbps.

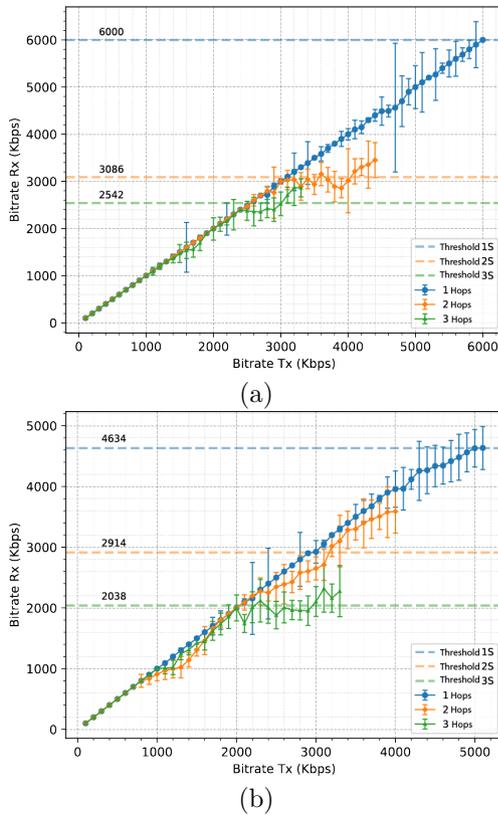


Figure 5. Behavior of the bitrate according to the number of hops. (a) With OLSR. (b) With BATMAN

Figure 5b illustrates the bitrate performance for the BATMAN protocol. For the 1-hop configuration (blue), the channel exhibits a favorable and nearly linear response, achieving 4.63 Mbps when traffic is transmitted at 5.1 Mbps. In the 2-hop scenario (orange), the response remains nearly linear up to 2.2 Mbps. Beyond this point, the channel begins to exhibit variability, reaching a maximum of 3.6 Mbps. Lastly, for the 3-hop configuration (green), the bitrate initially maintains linear behavior up to 1 Mbps. Beyond this threshold, the graph displays some variability, with the received bitrate stabilizing around 2 Mbps. However, when traffic exceeds 2 Mbps, the received bitrate fluctuates between 1.8 Mbps and 2.3 Mbps.

Based on the results, the threshold values corresponding to the maximum throughput achieved at each

hop were identified and are presented in Table 3.

The results presented in Figure 5 align with the findings in [31], which indicate that multi-hop flows tend to compete for access to the medium at each hop on their path to the destination node. Consequently, packets transmitted along longer routes are more likely to be discarded compared to those traveling shorter routes, thereby explaining the reduction in throughput observed at 2 and 3 hops.

Table 3. Maximum throughput achieved in the multi-hop network topology

Protocol	1 hop	2 hops	3 hops	Unit
OLSR	6.00	3.08	2.54	Mbps
BATMAN	4.63	2.91	2,038	

3.1. Selection of scenarios for network evaluation

Two scenarios were designed to evaluate the MANET network in conjunction with the mobile web application. Specifically, the area depicted in Figure 6 was designated for the tests, encompassing a 120-meter path within the CCTI-B facilities at the University of Cuenca.



Figure 6. Selected location for conducting the experiments within the CCTI-B

The experiments involved the transmission of voice traffic. The first scenario, illustrated in Figure 7, simulates the integration of a new node (node 4) into the MANET network. This node, serving as the data destination, moves along the defined path, while the remaining three nodes remain stationary, with node 2 functioning as the transmitter. Nodes 3 and 1 serve as intermediate hops to maintain communication when node 4 is positioned at the ends of the path.

The second scenario, illustrated in Figure 8, emulates communication between two moving cyclists (nodes 2 and 4), while nodes 1 and 3 remain stationary, serving as anchor stations to enable communication via hops. The fixed nodes are separated by a distance

of 60 meters, while the distance between the mobile nodes ranges from 30 to 90 meters. The results for each metric and scenario are detailed below.

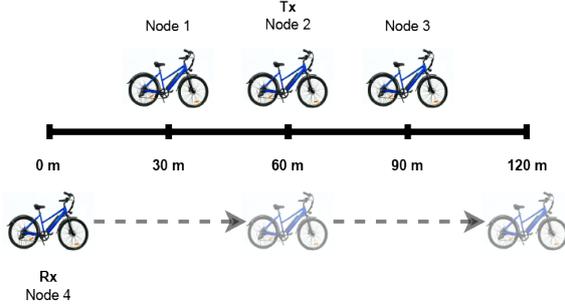


Figure 7. First scenario for the evaluation of the system and the mobile web application

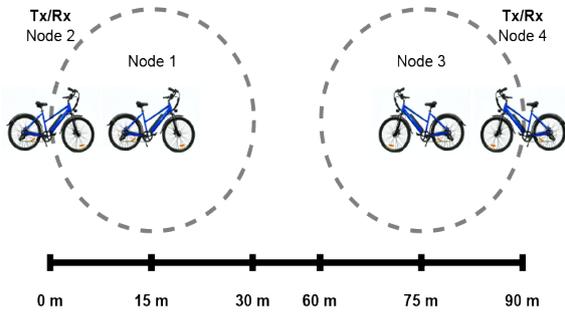


Figure 8. Second scenario for the evaluation of the system and the mobile web application

3.2. Delay

Figure 9 presents the results for the average delay observed with each protocol, analyzed with 95% confidence. Notably, the OLSR protocol exhibits a higher delay compared to BATMAN, with values of 49.5 ms versus 43.5 ms in the first scenario, and 20.9 ms versus 20.2 ms in the second scenario. This behavior can be attributed to the additional delay introduced by the exchange of signaling messages required for route establishment (e.g., Hello Interval, Orig Interval), as outlined in Table 1.

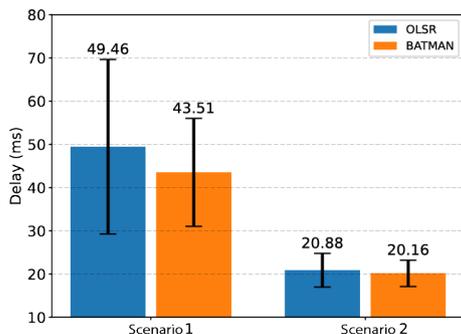


Figure 9. Average delay

3.3. PRR

Figure 10 presents the PRR percentages for scenarios 1 and 2. In the first scenario, the OLSR protocol achieves a PRR of approximately 96%, while the BATMAN protocol achieves 97%. Conversely, in the second scenario, OLSR records a PRR of 99%, outperforming BATMAN, which achieves 97%.

These results underscore the robust packet reception performance of both evaluated protocols. In the first scenario, both solutions achieve a reception rate exceeding 95%, while in the second scenario, this value rises to over 96%. Collectively, these findings highlight the effectiveness of both protocols in the evaluated scenarios.

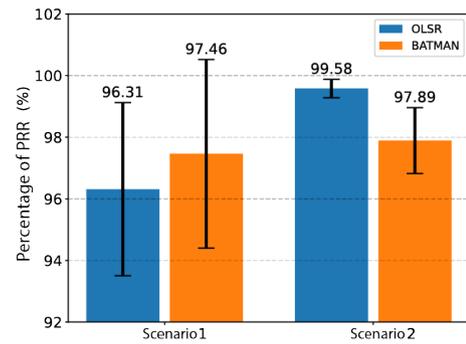


Figure 10. Average PRR

3.4. Throughput

Figure 11 illustrates the throughput results for each protocol. For the OLSR protocol, node 2 achieves a throughput of 68.21 Kbps, while node 4 records 65.7 Kbps. In the case of the BATMAN protocol, throughput values of 69.1 Kbps at node 2 and 65.82 Kbps at node 4 are observed.

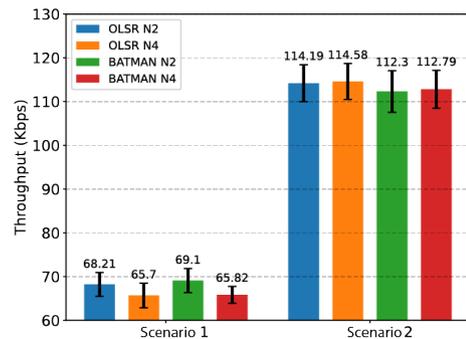


Figure 11. Average throughput

For scenario 2, the OLSR protocol achieves a throughput of 114.19 Kbps at node 2 and 114.58 Kbps at node 4. Similarly, for the BATMAN protocol, the observed throughput values are 112.3 Kbps at node 2 and 112.79 Kbps at node 4.

It is important to note that in the first scenario, only node 2 functioned as the transmitter and node 4 as the receiver, whereas in the second scenario, nodes 2 and 4 were configured for bidirectional communication. Consequently, higher throughput values are observed in the second scenario. Furthermore, these values remain within the multi-hop bandwidth capacity for both routing protocols.

3.5. Energy Consumption

Table 4 provides the energy measurements for each node in both scenarios. These measurements consider the maximum current consumed by the wireless interface and the 5 V operating voltage supplied by the USB (Universal Serial Bus) Type A port.

Table 4. Energy consumption at each node

Scenario	Protocol	Energy mAh			
		Node 1	Node 2	Node 3	Node 4
1	OLSR	206.75	201.5	176.6	205.53
	BATMAN	207.05	199.95	176.6	205.94
2	OLSR	79.22	72.57	79.42	76175
	BATMAN	87.57	72.93	88.3	72410

An analysis was performed to estimate the autonomy of the nodes based on their energy consumption during the experiments conducted in the two scenarios. The calculation utilized the 10 Ah capacity of the Eco move Electric Bikes battery [32] and the maximum current measurements for each node obtained using the INA219 current sensor [33]. Table 5 presents the estimated autonomy in hours.

Table 5. Estimation of autonomy time (h) for each node

Scenario	Protocol	Estimated autonomy time h			
		Node 1	Node 2	Node 3	Node 4
1	OLSR	24.18	24.18	28.31	24.32
	BATMAN	24.12	25.00	28.31	24.27
2	OLSR	31.55	34.44	31.47	32.81
	BATMAN	28.54	34.27	28.31	34.52

It is important to highlight that in some cases, the nodes' autonomy exceeds 24 hours. This can be attributed to the selection of test scenarios involving minimal motor usage, resulting in reduced battery consumption for the bicycle. Furthermore, during the experiments, only short-distance trips were conducted at speeds below 15 km/h.

4. Conclusions

This research proposes a MANET-based solution as an emergency communication system within the framework of sustainable mobility, specifically utilizing electric bicycles. A network comprising four mobile nodes and one primary fixed node was configured to conduct experiments, enabling the evaluation of proactive

routing protocols, OLSR and BATMAN, in multi-hop topologies under mobility conditions.

Based on the experimental results, the OLSR protocol demonstrated superior performance in terms of bandwidth in multi-hop network topologies. The maximum throughput achieved at one hop was 6 Mbps, decreasing progressively with the number of hops to 3.08 Mbps for two hops and 2.54 Mbps for three hops. Conversely, the BATMAN protocol exhibited a similar trend but achieved lower threshold values, with maximum throughputs of 4.63 Mbps for one hop, 2.91 Mbps for two hops, and 2.038 Mbps for three hops.

Additionally, the average delay values for scenario 1 remained below 50 ms, while for scenario 2, they were approximately 20 ms. These results indicate favorable communication performance for both protocols, aligning with the ITU-T G114 recommendations, which consider delays of up to 150 ms acceptable for real-time communication. Similarly, the PRR percentages demonstrated strong performance, with packet reception rates exceeding 96% in both scenarios. Lastly, the throughput results confirmed that this metric aligns with the channel's evaluated bandwidth capacity, ensuring reliable and seamless communication.

The analysis of the results for audio transmission, identified as the most critical case due to its stringent delay and PRR requirements, indicates that the OLSR protocol exhibited superior adaptability in the experiments conducted in this study. Moreover, the findings underscore the feasibility of leveraging emerging technologies such as MANET for the development of communication systems within the context of sustainable mobility.

Acknowledgments

The authors express their gratitude to the University of Cuenca for granting access to the Microgrid Laboratory at CCTI-B, providing the necessary equipment, and authorizing technical support from its staff during the experiments described in this article.

References

- [1] E. Salmeron-Manzano and F. Manzano-Agugliaro, "The electric bicycle: Worldwide research trends," *Energies*, vol. 11, no. 7, 2018. [Online]. Available: <https://doi.org/10.3390/en11071894>
- [2] K. Pangbourne, D. Stead, M. Mladenovic, and D. Milakis, *The Case of Mobility as a Service: A Critical Reflection on Challenges for Urban Transport and Mobility Governance*. United Kingdom: Emerald, 2018, pp. 33–48. [Online]. Available: <https://upsalesiana.ec/ing33ar3r2>

- [3] X. Xia, H. Jiang, and J. Wang, "Analysis of user satisfaction of shared bicycles based on sem," *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, no. 3, pp. 1587–1601, Mar 2022. [Online]. Available: <https://doi.org/10.1007/s12652-019-01422-y>
- [4] Z. Yang, J. Chen, J. Hu, Y. Shu, and P. Cheng, "Mobility modeling and data-driven closed-loop prediction in bike-sharing systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 12, pp. 4488–4499, 2019. [Online]. Available: <https://doi.org/10.1109/TITS.2018.2886456>
- [5] S. Shen, Z.-Q. Wei, L.-J. Sun, Y.-Q. Su, R.-C. Wang, and H.-M. Jiang, "The shared bicycle and its network—internet of shared bicycle (iosb): A review and survey," *Sensors*, vol. 18, no. 8, 2018. [Online]. Available: <https://doi.org/10.3390/s18082581>
- [6] F. Chen, K. Turoń, M. Kłos, W. Pamuła, G. Sierpiński, and P. Czech, "Fifth generation bike-sharing systems: examples from poland and china," *Scientific Journal of Silesian University of Technology. Series Transport*, vol. 99, pp. 05–13, 05 2018. [Online]. Available: <http://dx.doi.org/10.20858/sjsutst.2018.99.1>
- [7] T. Bielinski and A. Wazna, "New generation of bike sharing systems in china: Lessons for european cities," *Journal of Management and Financial Sciences*, no. 33, pp. 25–42, 2019. [Online]. Available: <https://doi.org/10.33119/JMFS.2018.33.2>
- [8] S. Yoo, S. Hong, Y. Park, A. Okuyama, Z. Zhang, Y. Yoshida, and S. Managi, "Danger, Respect, and Indifference: Bike-Sharing Choices in Shanghai and Tokyo using Latent Choice Models," MPRA Paper 108312, 2021. [Online]. Available: <https://upsalesiana.ec/ing33ar3r8>
- [9] M. Frikha, *Ad Hoc Networks: Routing, Qos and Optimization*. Wiley, 2013. [Online]. Available: <https://upsalesiana.ec/ing33ar3r9>
- [10] P. Astudillo Picon, C. Quidne Romero, S. Gonzalez Martinez, and I. Palacios Serrano, "Evaluación y comparación de códecs de video para el despliegue de un sistema de comunicación resiliente," *Revista Tecnológica ESPOL*, vol. 34, no. 3, pp. 12–30, 2022. [Online]. Available: <https://doi.org/10.37815/rte.v34n3.935>
- [11] J. Loo, J. Lloret, and J. Ortiz, *Mobile Ad Hoc Networks*. Taylor & Francis, 2011. [Online]. Available: <https://doi.org/10.1201/b11447>
- [12] K. Polshchikov, S. Lazarev, and E. Kiseleva, "Decision-making supporting algorithm for choosing the duration of the audio communication session in a mobile ad-hoc network," *Revista de la Universidad del Zulia*, vol. 10, no. 27, pp. 101–107, dic. 2019. [Online]. Available: <https://upsalesiana.ec/ing33ar3r12>
- [13] M. A. Al-Absi, A. A. Al-Absi, M. Sain, and H. Lee, "Moving ad hoc networks—a comparative study," *Sustainability*, vol. 13, no. 11, 2021. [Online]. Available: <https://doi.org/10.3390/su13116187>
- [14] A. Rosa, P. Á. Costa, and J. Leitão, "Generalizing wireless ad hoc routing for future edge applications," in *Mobile and Ubiquitous Systems: Computing, Networking and Services*, T. Hara and H. Yamaguchi, Eds. Cham: Springer International Publishing, 2022, pp. 264–279. [Online]. Available: https://doi.org/10.1007/978-3-030-94822-1_15
- [15] L. Reis, D. Macedo, and J. Nogueira, "Autoconfiguração de rotas em redes ad-hoc de vants," in *Anais do XXVII Workshop de Gerência e Operação de Redes e Serviços*. Porto Alegre, RS, Brasil: SBC, 2022, pp. 99–112. [Online]. Available: <https://doi.org/10.5753/wgrs.2022.223504>
- [16] D. G.C., A. Ladas, Y. A. Sambo, H. Pervaiz, C. Politis, and M. A. Imran, "An overview of post-disaster emergency communication systems in the future networks," *IEEE Wireless Communications*, vol. 26, no. 6, pp. 132–139, 2019. [Online]. Available: <https://doi.org/10.1109/MWC.2019.1800467>
- [17] F. A. León Mateo, M. d. R. Cruz Felipe, and E. T. Zambrano Solorzano, "Revisión de estudios sobre sistemas fanet y drones para emergencias o desastres naturales," *Serie Científica de la Universidad de las Ciencias Informáticas*, vol. 15, no. 4, pp. 41–56, 2022. [Online]. Available: <https://upsalesiana.ec/ing33ar3r17>
- [18] A. M. Soomro, M. F. Bin Fudzee, M. Husain, H. M. Saim, G. Zaman, A. Rahman, H. AlUbaidan, and M. Nabil, "Comparative review of routing protocols in manet for future research in disaster management," *Journal of Communications*, 2022. [Online]. Available: <https://doi.org/10.12720/jcm>
- [19] M. O. Olusanya and O. R. Vincent, "A manet-based emergency communication system for environmental hazards using opportunistic routing," in *2020 International Conference in Mathematics, Computer Engineering and Computer Science (ICMCECS)*, 2020, pp. 1–6. [Online]. Available: <https://doi.org/10.1109/ICMCECS47690.2020.240894>

- [20] A. Guillen-Perez, A.-M. Montoya, J.-C. Sanchez-Aarnoutse, and M.-D. Cano, "A comparative performance evaluation of routing protocols for flying ad-hoc networks in real conditions," *Applied Sciences*, vol. 11, no. 10, 2021. [Online]. Available: <https://doi.org/10.3390/app11104363>
- [21] Wardi, Dewiani, M. Baharuddin, S. Panggalo, and M. F. B. Gufran, "Performance of routing protocol olsr and batman in multi-hop and mesh ad hoc network on raspberry pi," *IOP Conference Series: Materials Science and Engineering*, vol. 875, no. 1, p. 012046, jun 2020. [Online]. Available: <https://dx.doi.org/10.1088/1757-899X/875/1/012046>
- [22] J. Yi and C. Poellabauer, "Real-time multicast for wireless multihop networks," *Computers & Electrical Engineering*, vol. 36, no. 2, pp. 313–327, 2010, wireless ad hoc, Sensor and Mesh Networks. [Online]. Available: <https://doi.org/10.1016/j.compeleceng.2009.03.009>
- [23] Z. haitao, Z. yuting, Z. hongbo, and L. dapeng, "Resource management in vehicular ad hoc networks: Multi-parameter fuzzy optimization scheme," *Procedia Computer Science*, vol. 129, pp. 443–448, 2018, 2017 INTERNATIONAL CONFERENCE ON IDENTIFICATION, INFORMATION AND KNOWLEDGE IN THE INTERNET OF THINGS. [Online]. Available: <https://doi.org/10.1016/j.procs.2018.03.022>
- [24] M. Elaryh Makki Dafalla, R. A. Mokhtar, R. A. Saeed, H. Alhumyani, S. Abdel-Khalek, and M. Khayyat, "An optimized link state routing protocol for real time application over vehicular ad hoc network," *Alexandria Engineering Journal*, vol. 61, no. 6, pp. 4541–4556, 2022. [Online]. Available: <https://doi.org/10.1016/j.aej.2021.10.013>
- [25] K. A. Polshchikov, S. A. Lazarev, E. D. Kiseleva, E. M. Mamatov, and E. V. Ilin-skaya, "Audio communication quality provision in a self-organizing network," *Procedia Environmental Science, Engineering and Management*, vol. 9, pp. 509–515, 2022. [Online]. Available: <https://upsalesiana.ec/ing33ar3r25>
- [26] M. Hosseinzadeh, S. Ali, A. H. Mohammed, J. Lansky, S. Mildeova, M. S. Yousefpoor, E. Yousefpoor, O. Hassan Ahmed, A. M. Rahmani, and A. Mehmood, "An energy-aware routing scheme based on a virtual relay tunnel in flying ad hoc networks," *Alexandria Engineering Journal*, vol. 91, pp. 249–260, 2024. [Online]. Available: <https://doi.org/10.1016/j.aej.2024.02.006>
- [27] R. Prasad P and Shivashankar, "Enhanced energy efficient secure routing protocol for mobile ad-hoc network," *Global Transitions Proceedings*, vol. 3, no. 2, pp. 412–423, 2022, global Transitions 2019. [Online]. Available: <https://doi.org/10.1016/j.gltp.2021.10.001>
- [28] M. Arun and R. Jayanthi, "An adaptive congestion and energy aware multipath routing scheme for mobile ad-hoc networks through stable link prediction," *Measurement: Sensors*, vol. 30, p. 100926, 2023. [Online]. Available: <https://doi.org/10.1016/j.measen.2023.100926>
- [29] Iperf.fr. (2023) iperf - the tcp, udp and setp network bandwidth measurement tool. [Online]. Available: <https://upsalesiana.ec/ing33ar3r29>
- [30] FFmpeg. (2023) A complete, cross-platform solution to record, convert and stream audio and video. [Online]. Available: <https://upsalesiana.ec/ing33ar3r30>
- [31] W. E. Castellanos Hernández, "Quality of service routing and mechanisms for improving video streaming over mobile wireless ad hoc networks," Ph.D. dissertation, Universitat Politècnica de Valencia, 2015. [Online]. Available: <http://dx.doi.org/10.4995/Thesis/10251/53238>
- [32] Ecomove. (2023) Tiv - ecomove. [Online]. Available: <https://upsalesiana.ec/ing33ar3r32>
- [33] Adafruit. (2023) Adafruit ina219 current sensor breakout. [Online]. Available: <https://upsalesiana.ec/ing33ar3r33>



PROTOTYPE OF A REMOTELY CONTROLLED MULTICHANNEL SURFACE MUSCLE STIMULATOR

PROTOTIPO DE ESTIMULADOR MUSCULAR SUPERFICIAL MULTICANAL CONTROLADO REMOTAMENTE

P. Silverio-Cevallos^{1,*} , J. Maita Cajamarca¹ ,
D. A. Molina-Vidal² , C. J. Tierra-Criollo² , P. Cevallos-Larrea¹ 

Received: 13-11-2023, Received after review: 29-05-2024, Accepted: 05-11-2024, Published: 01-01-2025

Abstract

Multichannel Functional Electrical Stimulation (FES) technology is widely employed in artificial motor control research. This study presents the design and evaluation of a four-channel, remotely controlled surface electrical muscle stimulator prototype. The prototype introduces a modern alternative for the control block, employing a Wi-Fi-enabled solution based on the ESP32 microcontroller. This controller enables remote configuration of activation sequences for individual channels and supports extensive customization of parameters for a biphasic waveform stimulus. The current signal is demultiplexed into four outputs. Additionally, this study provides a detailed functional evaluation of the amplification stage and examines the load-dependent limitations of the output current magnitude. Preliminary experimental testing demonstrates the prototype's ability to generate controlled stimulation sequences in hand muscles. The prototype's functional and experimental performance suggests its potential application in artificial motor control research.

Keywords: Multichannel Functional Electrical Stimulator, Muscle Stimulator, Artificial Motor Control

Resumen

La tecnología de Estimulación Eléctrica Funcional (EEF) multicanal se utiliza actualmente en la investigación del control motor artificial. Este trabajo describe el diseño y evaluación de un prototipo de estimulador eléctrico muscular de cuatro canales controlado remotamente. El prototipo propone una alternativa moderna para el bloque de control, utilizando el microcontrolador Wi-Fi/ESP32. Este permite una secuencia de activación de canales configurable de manera remota y una extensiva configuración de los parámetros de un estímulo en forma de onda bifásica. La señal de corriente se demultiplexa en cuatro salidas. Este estudio también contribuye detallando la evaluación funcional de la etapa de amplificación y estableciendo la dependencia de la magnitud de la carga en los límites de la corriente de salida. La prueba experimental preliminar demuestra la capacidad del prototipo para generar secuencias de estimulación controladas en los músculos de la mano. El desempeño funcional y experimental del prototipo sugiere su potencial uso para investigaciones del control motor artificial.

Palabras clave: Estimulador Eléctrico Funcional Multicanal, Estimulador Muscular, Control Motor Artificial

^{1,*}Biomedical Engineering Research Group - GIIB, Universidad Politécnica Salesiana, Ecuador.

Corresponding author ✉: psilverio@est.ups.edu.ec.

² Biomedical Engineering Program, Alberto Luiz Coimbra Institute for Graduate Studies and Research in Engineering (Coppe), Federal University of Rio de Janeiro, Brazil.

Suggested citation: Silverio-Cevallos, P., Maita Cajamarca, J., Molina-Vidal, D. A., Tierra-Criollo, C. J. and Cevallos-Larrea, P. "Prototype of a remotely controlled multichannel surface muscle stimulator," *Ingenius, Revista de Ciencia y Tecnología*, N.º 33, pp. 38-48, 2025, DOI: <https://doi.org/10.17163/ings.n33.2025.04>.

1. Introduction

Neurodegenerative disorders, such as Parkinson’s disease (PD), spinal muscular atrophy, and amyotrophic lateral sclerosis, among others [1–3], have a profound impact on the nervous system, frequently affecting motor functions. Common symptoms associated with these motor disorders include difficulty initiating and coordinating smooth muscular movements, inhibition of involuntary movements, challenges with postural adjustment, progressive limb muscle weakness, and muscle atrophy [4–6].

Electrical stimulation therapy plays a pivotal role as a non-pharmacological treatment for motor disorders associated with neurodegenerative diseases. Common non-invasive techniques include Transcutaneous Electrical Nerve Stimulation (TENS) and Functional Electrical Stimulation (FES). TENS primarily targets afferent nerve fibers to mitigate muscle atrophy, alleviate pain, enhance muscle strength, and support functional movement therapy [7]. In contrast, FES stimulates motor nerves to induce contractions in weak or paralyzed muscles. This technique is particularly effective for patients with motor impairments, such as those experiencing paralysis or severe muscle weakness [8].

Artificial motor control through FES is an assistive strategy designed to achieve functional and intentional movements by inducing controlled contractions in targeted muscle groups [9]. The therapeutic potential of this technique has been extensively studied in various conditions, including Parkinson’s disease [10], paraplegia [11] and neuroprosthetics [12].

Research conducted by Qi Wu et al. [13] and Masdar et al. [14] demonstrated the efficacy of electrical stimulation in restoring and maintaining muscle activity in paralyzed patients with spinal cord injuries and related neural deficits. Furthermore, studies by Hai-Peng Wang et al. [15] and Keller T. [16] highlighted the use of electrical stimulation to enhance motor control and support motor function training in stroke patients.

Despite the availability of various commercial electrical stimulation technologies, experimental paradigms in artificial motor control often require stimulators with capabilities that surpass those offered by standard TENS and FES technologies [17, 18]. For instance, advanced features such as multichannel stimulation with remotely programmable output patterns and customizable stimulus parameters are critical in this context [19].

However, detailed accounts of such advanced electrical stimulation prototypes remain limited. One notable example is the multichannel programmable stimulator prototype developed by Qi Xu et al. [20].

Similarly, Hai-Peng Wang et al. [15] proposed a FES stimulator capable of multiplexing signals from an amplifier circuit across multiple outputs with pro-

grammable stimuli. Despite these advancements, current reports on multichannel electrical stimulation technology for complex motor control exhibit significant limitations [15], [20].

First, many prior studies rely on electronic control blocks that are difficult to procure or replicate due to inaccessible developer tools and documentation [15], [19]. For instance, the prototypes described in [15], [20] employ outdated controller technologies.

Second, these reports often lack comprehensive descriptions of performance evaluations and the limitations of signal amplification and current source circuitry, impeding reproducibility and validation efforts.

To address the need for contrasting and replicable research in advanced motor control, it is essential to explore modern and easily replicable technologies for electrical stimulation. This study aims to design and evaluate a prototype multichannel, wireless surface electrical stimulator for FES. The design specifications include remote control functionality via a smartphone and the use of widely available electronic components with extensive development resources to facilitate replication.

Additionally, the prototype is capable of generating programmable sequences of multiplexed rectangular biphasic signals across four isolated channels. The controller block is implemented using the ESP32 wireless microcontroller, a widely adopted platform known for its large support community, versatility, and scalability [21–24].

Furthermore, a preliminary experimental test was conducted to assess the prototype’s ability to generate sequential, programmable muscle contractions in hand muscles, captured using a sensorized glove equipped with accelerometers.

2. Materials and methods

2.1. Design Methodology

Figure 1 illustrates the general architecture of the Multichannel Surface Electrical Stimulator (MSSES). The system consists of two primary components: (i) a hardware module that generates biphasic waveform current stimuli across four asynchronously activated channels, and (ii) a software module, implemented as a smartphone application, which allows users to configure stimulation parameters, including magnitude, total period, inter-stimulus interval, and stimulation sequences across output channels. Communication between the hardware and software modules is facilitated through a wireless (Wi-Fi) connection.

2.2. Hardware

The hardware architecture consists of three main blocks: sourcing, control, and current stimulus genera-

tion, as depicted in Figure 2.

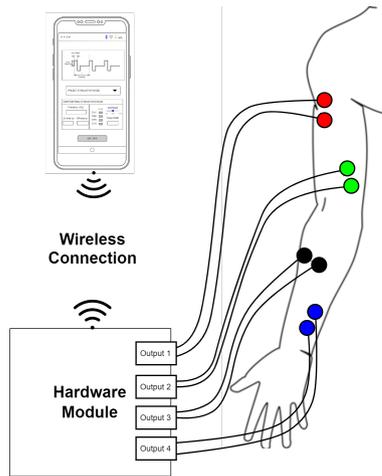


Figure 1. General architecture of the MSES

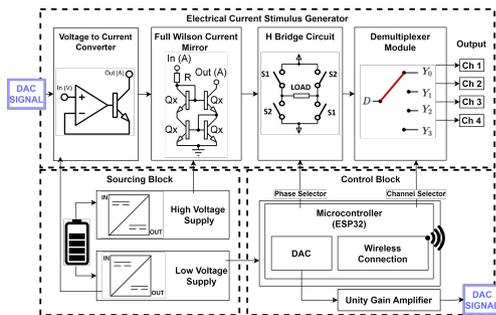


Figure 2. Hardware architecture of MSES

The Sourcing Block is powered by a 5 VDC battery, generating two isolated voltage levels. A low voltage level of ± 5 VDC is provided by an isolated DC-DC converter (model THM 10-0521WI) to power the digital circuits within the control block and the analog signal conditioning circuit in the initial stage of the current stimulus generator. Additionally, a high voltage level of ± 60 VDC is generated using an isolated

DC-DC converter (model R05-100B) to supply the current stimulus generator.

The Control Block is managed by an ESP32 microcontroller (Ten silica Xtensa, 32-bit, LX6 processor), featuring integrated wireless communication capabilities. The firmware algorithm processes incoming commands—such as start, stop, update stimulus, and channel sequence—as well as stimulus parameters, including anodic and cathodic current periods, magnitude, and inter-channel intervals, from the remote application. The control block sets the low-level stimulus amplitude using an 8-bit digital-to-analog converter (DAC) connected to a unity-gain amplifier circuit, providing a DAC output range of 0 to 3.3 VDC. Additionally, it performs two critical functions within the current stimulus generator block: reversing the amplified magnitude of the electrical pulse to produce a biphasic stimulus and demultiplexing the stimulus to a designated output channel.

The Electrical Current Stimulus Generator comprises four main stages. The first stage converts the low-level voltage from the DAC signal into a high-voltage-driven current signal (± 60 VDC), adhering to the recommendations outlined in [15]. Specifically, the DAC signal is fed into a voltage-to-current converter circuit, commonly referred to as a transconductance amplifier. The resulting current signal drives two current amplification circuits (Wilson Current Mirror - WCM), each powered by the levels HV+ and HV-, creating a constant current flow through OUT+ and OUT- when a load is connected, as depicted in Figure 3. In the WCM circuit, using resistor values ($R+$ and $R-$) lower than $1\text{ k}\Omega$ results in signal degradation at the output, particularly for $\approx 1\text{ k}\Omega$. This study adopted resistor values of $2.4\text{ k}\Omega$ for $R+$ and $R-$, which demonstrated the lowest noise levels at the output and minimized voltage drop across VCE in transistors Q1 and Q2. The specific values of $R+$ and $R-$ also influence the maximum voltage at OUT+ and OUT-, consequently limiting the maximum current output [15].

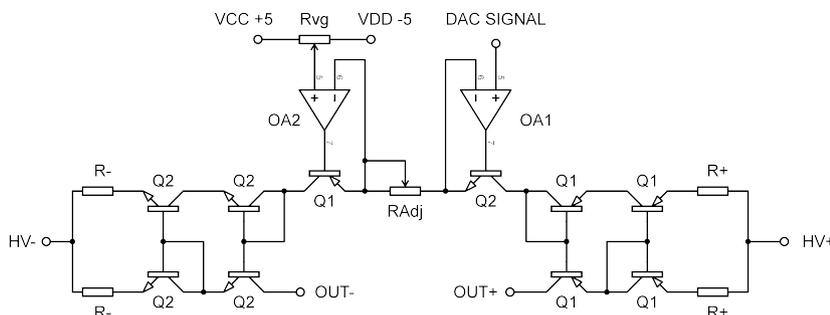


Figure 3. Electrical Schematic of the Voltage-to-Current Converter (V-to-C) Circuit and the Wilson Current Mirror (WCM) Circuits. The V-to-C circuit uses Op Amps TLC2252 (OA1 and OA2) and a variable resistor (RAdj) to adjust the working current range. The WCM employs NPN transistors (2N6517, Q2) with resistor $R-$ to amplify the signal to HV-, and PNP transistors (2N6520, Q1) with resistor $R+$ to achieve HV+

In the second stage, the output terminal of the WCM is connected to a phase inverter circuit featuring a programmable four-switch H-bridge topology controlled by the control block. This circuit reverses the direction of current flow through the load to generate a biphasic waveform or disables stimulus transmission entirely. The hardware can be configured to produce a square waveform with specified durations for the anodic cycle (anodic current, S1 switches ON), cathodic cycle (cathodic current, S2 switches ON), and the interval between these cycles. Finally, the biphasic signal passes through a selector circuit (demultiplexer) that, based on the configuration of the control block, directs the stimulus to one of the four available outputs.

2.3. Software

This project utilized the Modular platform to develop a smartphone-based user interface application, as depicted in Figure 4 [25]. The application enables users to adjust stimulus parameters (Figure 4a), including stimulus magnitude (I+, I-), total period (T), anodic current period (Tp, positive), cathodic current period (Tn, negative), and two unstimulated periods: Tc1 (between Tp and Tn) and Tc2 (following Tn).

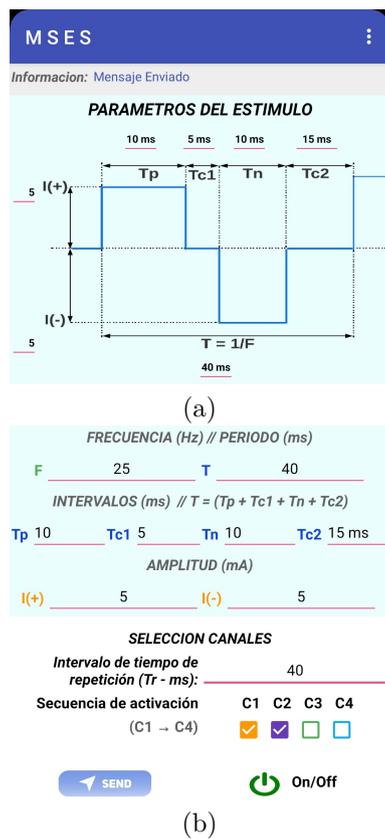


Figure 4. User Interface Application. (a) Waveform panel displaying stimulus parameters. (b) Controls for inputting stimulus parameters and managing the stimulation protocol

When stimulation is applied to a specific channel, the stimulus is delivered repeatedly to that output over a user-configurable period, Tr, ensuring consistent mechanical contractions that may not be achievable with a single square stimulus pulse. The user interface supports real-time updates of stimulation parameters in the hardware and manages the application of stimuli (Figure 4b).

During multichannel stimulation, the configured stimulus is sequentially directed to the enabled channels in ascending order (i.e., from channel 1 to channel 4).

2.4. Performance Tests

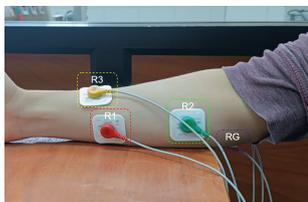
The evaluation of the system's performance involves determining the operational limits of the stimulus parameters and the multichannel stimulation paradigm. The first test assessed the stimulus magnitude by activating a single output (channel 1) and varying both the stimulus magnitude and the resistive load values (1 kΩ, 5 kΩ, and 10 kΩ). During this test, the DAC output was set to nine fixed values within its dynamic range, enabling the derivation of an equation relating the digital values configured in the DAC to the output current levels. The second test measured the stability of the electrical current and the rising and falling times of the square wave stimulus under varying load conditions. The test used a fixed current level of 5 mA, with period values for Tp, Tn, Tc1, Tc2, and total period (T) set to 25 ms, 25 ms, 10 ms, 40 ms, and 100 ms, respectively (F = 10 Hz). Resistive load values of 1, 3.3, 5.6, 10, 12, and 20 kΩ were applied. The final test evaluated the system's capability to sequentially redirect the configured stimulus across multiple channels, following the multichannel stimulation paradigm. This test simultaneously measured two channels (channels 1 and 2), using stimulus magnitudes of 2.5 mA and 5 mA, periods Tp and Tn of 25 ms and 40 ms, respectively, and a Tr period equal to the total period T (one stimulus per channel).

2.5. Preliminary application test

This preliminary application test served as an initial evaluation of the proposed technology, without extensive testing on healthy individuals or patients with neurodegenerative diseases. The primary objective was to assess the system's ability to generate controlled electrical current stimuli across multiple channels, inducing intentional finger contractions in a predetermined sequence. The experiment was conducted in the Laboratory of the Biomedical Engineering Research Group (GIIB-UPS) with two participants, both authors of this study. Both participants reported being in good health, with no history of muscular or neurological disorders, cardiac conditions, or pacemaker

use. This test adhered to the ethical principles outlined in the Declaration of Helsinki [26], and informed consent was obtained from both participants. Three stimulation regions (R1, R2, R3) on the forearm and a ground region (RG) on the olecranon were identified for electrode placement (Figure 5a). This configuration followed a previously established protocol [27] to elicit contractions in the index, middle, and ring + little fingers, corresponding to stimulation in R1, R2, and R3, respectively.

Before initiating multichannel stimulation, the stimulus magnitude was determined to elicit visible but painless muscular contractions. To achieve this, single-channel stimulation was applied, gradually increasing the current level from 0 mA until a visible contraction was observed, ensuring the absence of pain for the participant. The parameters selected for the multichannel test were as follows: total period (T) of 20 ms, anodic (Tp) and cathodic (Tn) phases of 200 μ s, an inter-phase interval (Tc1) of 100 μ s, a repetition interval (Tr) of 5 s, and a stimulus magnitude of approximately 5 mA. These stimulation levels align with those used in previous studies [28].



(a)



(b)

Figure 5. (a) Electrode placement regions for forearm stimulation (R1, R2, R3) and reference in RG., (b) Sensing glove with MPU6050 sensors attached to index, middle, ring, and little fingers

For multichannel stimulation, three channels (C1, C2, C3) and two stimulation modes were utilized. In the first mode, channels C1, C2, and C3 were connected to regions R1, R2, and R3, respectively. In the

second mode, the connections were reconfigured to R2, R3, and R1, respectively. In both modes, the stimulator was programmed with a sequential stimulation pattern of $C1 \rightarrow C2 \rightarrow C3$.

A sensing glove was developed to objectively measure finger movement in response to each stimulus. This glove incorporates four acceleration sensors (MPU6050), each attached to the index, middle, ring, and little fingers (Figure 5b). The sensors communicate with an AT mega 328 microcontroller (Arduino Nano) via the I2C protocol, and the recorded data are stored in a digital .txt file using serial communication. The sensor data facilitate the calculation of the rotation angle (pitch) for each finger as it flexes. Before stimulation, the participant was instructed to maintain their hand in a natural, relaxed position (rest), during which the initial mean rotation angles were recorded. Consequently, the sensor data are expressed as values relative to the sensors initial positions.

3. Results and discussion

3.1. Performance Indicators

The DAC's output varied linearly within a range of -0.08 to 2.93 V for input values between 0 and 255 digital units. To achieve a 10 mA output in the electric current generator block from the maximum DAC output voltage, R_{Adj} in the V-to-C circuit (Figure 3) was set according to the equation: $R_{Adj} = V_{DAC_{max}}/I_{max}$, that is $2.93V/10mA = 293\Omega$. R_{Adj} adjusts the current level at the output of the V-to-C circuit, which is subsequently amplified through the WCM circuit (Figure 3). Figure 6a illustrates a directly proportional relation between the voltage across the load and the DAC levels. The output voltage increases with the load value to maintain a fixed current level at the output. However, as the load magnitude increases, both the output voltage and the current level reach saturation. This behavior is attributed to the maximum voltage available at the OUT+/- terminals during the performance test, which reached a maximum value of 77.6 VDC. It is important to note that during experimental tests, the DC-DC converters were set to provide +/-64 V for +/-HV. Figure 6b demonstrates an approximately linear relationship between the output current and the DAC control variable for three resistive loads (1, 5, and 10 k Ω). A maximum current of 7.63 mA was achieved for the 10 k Ω load, consistent with the saturation explanation provided earlier. For the 5 k Ω load, the linear trend was calculated using least-squares regression to determine the output/input relationship. This analysis yielded the equation: $I_{out}(mA) = (0.038 \times \text{digital value}) - 0.0819$, is integrated into the firmware algorithm to convert the stimulus magnitude, expressed in units of electric current, into digital DAC values: $\text{digital value} = (1000 \times I_{out} + 81.9)/38.8$.

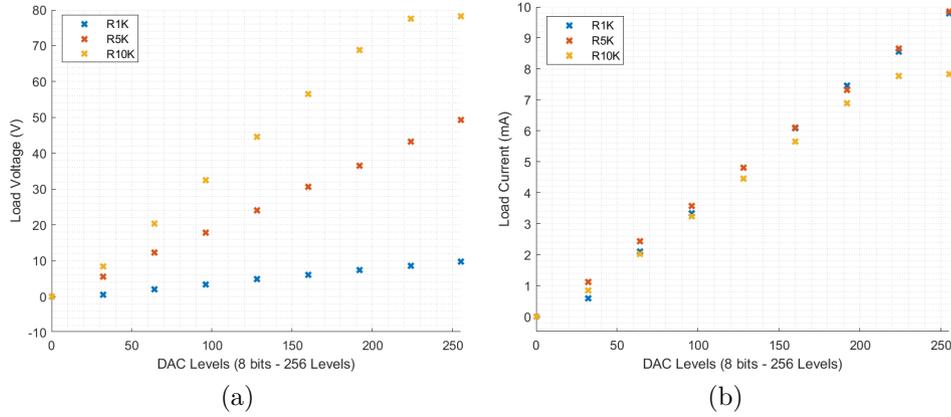


Figure 6. Voltage and current measurements for a single stimulus applied to resistive loads of 1, 5, 10 $k\Omega$, within the range of the DAC control variable. (a) Load output voltage vs DAC Levels, (b) Load output current vs DAC Levels

Figure 7a illustrates the waveform of the output current signal for load resistances ranging from 1 $k\Omega$ to $k\Omega$. Overall, the measured magnitude of the biphasic stimulus remains stable, with a mean value of 4.38 mA and a standard deviation of $\pm 618 \mu A$ (12.37 % relative to the stimulus magnitude). The maximum variation (14.5 %) was observed at a load of 20 $k\Omega$. Figure 7b depicts the signal transition time measurements for changes between stimulus magnitudes of 10

% to 90 % and vice versa. The average rise time was 10.1 μs , with minimum and maximum times of 1.6 μs and 11.7 μs , respectively. Some non-linearities in signal magnitude, such as a peak at the start of the transition, were noted as the load resistance decreased. Conversely, the average fall time was 10.9 μs , with minimum and maximum times of 0.4 μs and 11.3 μs , respectively. In general, both rise and fall times increased slightly as the load resistance increased.

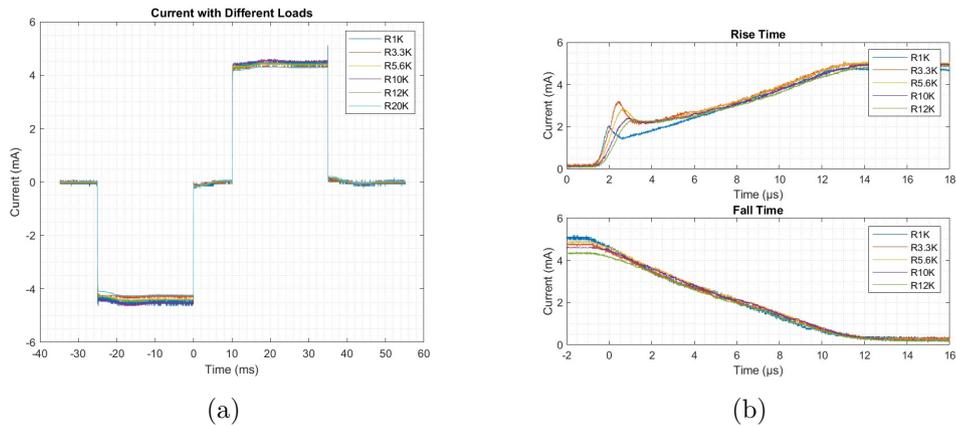


Figure 7. Stimulus waveform parameters for different load resistances: (a) Stability of the biphasic stimulus waveform current, (b) Rise and fall times for a stimulus magnitude of 5 mA

Names should be abbreviated using initials only. The amplitudes and periods generated by the MSES closely matched those configured in the user interface, as shown in Figure 8. This figure illustrates a sequence of stimuli generated on channels one and two, with variations in some stimulus parameters. In Figure 8a, the values for T_p , T_n , and magnitude were 40 ms, 40

ms, and 2.72 mA, respectively, while in Figure 8b, these values were 25 ms, 25 ms, and 5.26 mA. The total period (T), which was set equal to T_r for this test, was 100 ms. Additionally, Figure 8 demonstrates the absence of interchannel interference during sequential stimulation on channels one and two.

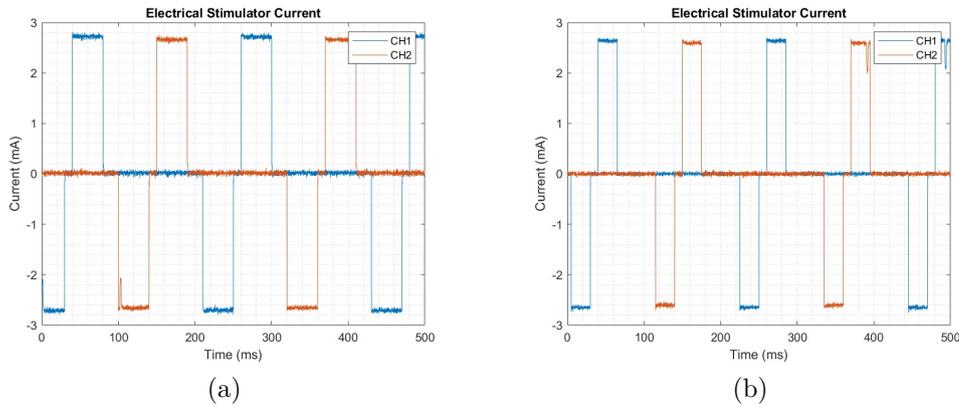


Figure 8. Output of two synchronously applied channels. (a) signal with T_p and T_n of 40 ms and (b) signal T_p and T_n of 25 ms.

3.2. Application test

Finger contraction and relaxation events, along with their relationship to the two proposed stimulation sequences, were analyzed using the signals recorded by the sensorized glove for participants #1 and #2 (Fi-

gure 9). Overall, the sensor data demonstrated that specific finger movement patterns are primarily influenced by the stimulated region, with less influence from the stimulation sequence ($R1 \rightarrow R2 \rightarrow R3$ or $R2 \rightarrow R3 \rightarrow R1$).

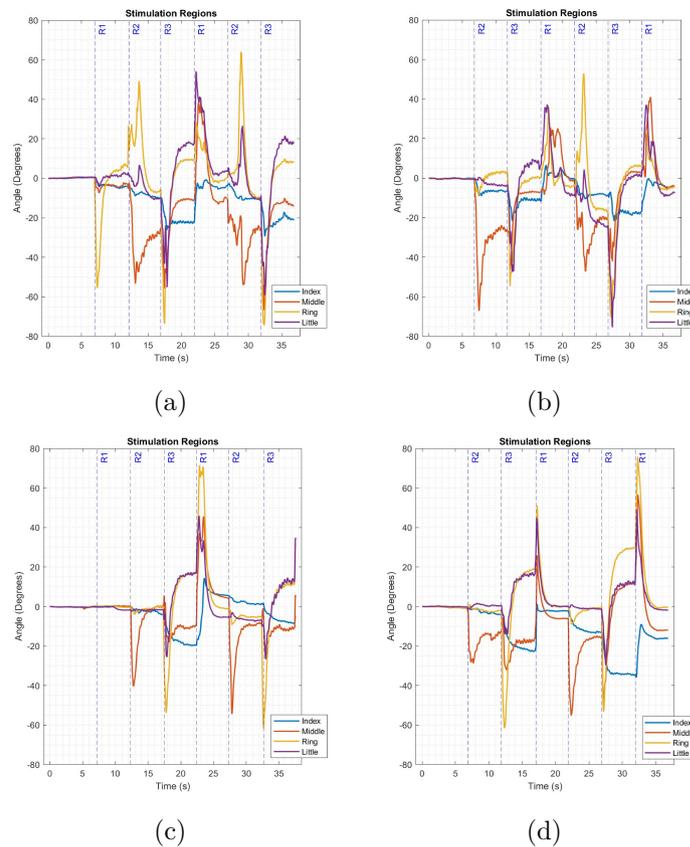


Figure 9. Rotation angles obtained from acceleration sensors attached to the sensitized glove. The movement patterns correspond to sequences $R1 \rightarrow R2 \rightarrow R3$ for (a) and (c), and $R2 \rightarrow R3 \rightarrow R1$ for (b) and (d), and participant #1 with (a) and (b) and #2 with (c) and (d)

For example, in participant #1, stimulation of R2 predominantly caused contraction of the middle finger and relaxation tendencies in the ring and pinky fingers. This pattern was observed in both repetitions of the sequence $R1 \rightarrow R2 \rightarrow R3$ (Figure 9a) and once in the sequence $R2 \rightarrow R3 \rightarrow R1$ (Figure 9b). Stimulation sequences beginning at R2 (from the resting state) did not generate signals associated with ring and pinky finger relaxation. Additionally, stimulation of R3 induced contraction of the ring, middle, and pinky fingers for both sequences in participant #1. Conversely, stimulation of R1 generally exhibited a relaxation effect, particularly when any fingers were previously contracted. This effect was evident in the sequence $R2 \rightarrow R3 \rightarrow R1$ (Figure 9b) and during the second R1 stimulation in the sequence $R1 \rightarrow R2 \rightarrow R3$ (Figure 9b). When stimulation began at R1 (after the resting state, Figure 9a), a contraction movement was observed in the ring finger. The contraction patterns observed in participant #2 were similar to those noted in participant #1.

In summary, the stimulation sequence $R1 \rightarrow R2 \rightarrow R3$ (Figure 9c) elicited the following pattern of movements: no finger contraction (R1) \rightarrow predominant contraction of the middle finger (R2) \rightarrow predominant contraction of the ring finger, with less pronounced contractions of the middle and pinky fingers (R3). Conversely, the sequence $R2 \rightarrow R3 \rightarrow R1$ (Figure 9d) produced the following movement pattern: predominant contraction of the middle finger (R2) \rightarrow predominant contraction of the ring finger, along with less pronounced contractions of the middle and pinky fingers (R3) \rightarrow no contraction or relaxation of previously contracted fingers (R1). Predominant contraction events for both participants and stimulation sequences are summarized in Table 1.

Table 1. Summary of finger contractions or relaxing events for two stimulation sequences in both participants (P#1, P#2).

Seq	R1 \rightarrow R2 \rightarrow R3			R2 \rightarrow R3 \rightarrow R1		
P#1	ring	middle	ring	middle	ring	relaxing
	---	---	mean/ little	---	mean/ little	---
P#2	relaxing	middle	ring	middle	ring	relaxing
	---	---	mean/ little	---	mean/ little	---

4. Conclusions

This study presents a prototype for a Functional Electrical Stimulation (FES) multichannel system capable of delivering programmable sequences of multiplexed rectangular biphasic signals across four isolated channels, with operational control via a smartphone.

The proposed prototype offers several technological and documentation advancements compared to prior research on similar electrical stimulation technologies. First, the current design employs the widely

accessible, cost-effective, and well-supported ESP32 wireless microcontroller. This modern controller simplifies the stimulator's electronic control block, addressing limitations in earlier designs that relied on outdated controllers, as noted in [15], [20]. This modification provides a replicable alternative for the control block of stimulation technologies, potentially facilitating further research in artificial motor control. Second, this study provides detailed performance data for the circuitry within the stimulus generation block, a feature not addressed in prior designs of multichannel electrical stimulators [15], [20]. Specifically, the current output exhibited a strong dependence on the adjustment of resistors R_{Adj} and R_{Vg} (Figure 3). While the circuit effectively generates constant current stimuli, the maximum current level and the dynamic range of the amplifier stage are constrained by increases in the load magnitude connected to the output. Third, this prototype introduces a scalable multiplexing scheme utilizing a combination of optocoupler and triac per channel. This topology enables straightforward replication to expand the number of channels as needed.

Additionally, the preliminary tests demonstrated the system's capability to generate programmable sequences of controlled muscle contractions.

The results suggest that the prototype is well-suited for integration into extended experimental protocols for multichannel sequential muscle stimulation. Future work will focus on developing a miniaturized, embedded version of the prototype in the form of a handheld device equipped with an accelerometer. This enhanced iteration will facilitate broader experimental applications of multichannel sequential muscle stimulation, including studies on its impact on conditions such as Parkinson's disease and stroke.

References

- [1] R. N. L. Lamptey, B. Chaulagain, R. Trivedi, A. Gothwal, B. Layek, and J. Singh, "A review of the common neurodegenerative disorders: Current therapeutic approaches and the potential role of nanotherapeutics," *International Journal of Molecular Sciences*, vol. 23, no. 3, 2022. [Online]. Available: <https://doi.org/10.3390/ijms23031851>
- [2] C. H. Zapata-Zapata, E. Franco-Dáger, J. M. Solano-Atehortúa, and L. F. Ahunca-Velásquez, "Esclerosis lateral amiotrófica: actualización," *Revista Médica Iatreia*, vol. 29, no. 2, pp. 194–205, 2016. [Online]. Available: <https://doi.org/10.17533/udea.iatreia.v29n2a08>
- [3] N. Cardona, S. J. Ocampo, J. M. Estrada, M. I. Mojica, and G. L. Porras, "Caracterización clínica y funcional de pacientes con atrofia muscular espinal en el centro-occidente colombiano,"

- Revista del Instituto Nacional de Salud Biomedica*, vol. 42, no. 1, pp. 59–99, 2022. [Online]. Available: <https://doi.org/10.7705/biomedica.6178>
- [4] L. Meng, M. Jin, X. Zhu, and D. Ming, “Peripheral electrical stimulation for parkinsonian tremor: A systematic review,” *Frontiers in Aging Neuroscience*, vol. 14, 2022. [Online]. Available: <https://doi.org/10.3389/fnagi.2022.795454>
- [5] O. Hardiman, A. Al-Chalabi, A. Chio, E. M. Corr, G. Logroscino, W. Robberecht, P. J. Shaw, Z. Simmons, and L. H. van den Berg, “Amyotrophic lateral sclerosis,” *Nature Reviews Disease Primers*, vol. 3, no. 1, p. 17071, Oct 2017. [Online]. Available: <https://doi.org/10.1038/nrdp.2017.71>
- [6] M. Gobbo, S. Lazzarini, L. Vacchi, P. Gaffurini, L. Bissolotti, A. Padovani, and M. Filosto, “Exercise combined with electrotherapy enhances motor function in an adolescent with spinal muscular atrophy type iii,” *Case Reports in Neurological Medicine*, vol. 2019, no. 1, p. 4839793, 2019. [Online]. Available: <https://doi.org/10.1155/2019/4839793>
- [7] A. Wright and K. A. Sluka, “Nonpharmacological treatments for musculoskeletal pain,” *The Clinical Journal of Pain*, vol. 17, no. 1, 2001. [Online]. Available: <https://doi.org/10.1097/00002508-200103000-00006>
- [8] D. B. Popovic, “Advances in functional electrical stimulation (fes),” *Journal of Electromyography and Kinesiology*, vol. 24, no. 6, pp. 795–802, 2014. [Online]. Available: <https://doi.org/10.1016/j.jelekin.2014.09.008>
- [9] C. Marquez-Chin and M. R. Popovic, “Functional electrical stimulation therapy for restoration of motor function after spinal cord injury and stroke: a review,” *BioMedical Engineering OnLine*, vol. 19, no. 1, p. 34, May 2020. [Online]. Available: <https://doi.org/10.1186/s12938-020-00773-4>
- [10] M. Djurić-Jovičić, S. Radovanović, I. Petrović, C. Azevedo, G. Mann, and M. Popović, “The impact of functional electrical stimulation (fes) on freezing of gait (fog) in patients with parkinson’s disease,” *Clinical Neurophysiology*, vol. 124, no. 7, p. e11, 2013. [Online]. Available: <https://doi.org/10.1016/j.clinph.2012.12.019>
- [11] F. Naz, D. Hussain, H. Ali, Q. Raza, and F. Siddique, “Effectiveness of functional electrical stimulation machine in managing neurological diseases - a retrospective study,” *Pakistan Journal of Medical Sciences*, vol. 40, no. 2(ICON), Dec. 2023. [Online]. Available: [https://doi.org/10.12669/pjms.40.2\(ICON\).8966](https://doi.org/10.12669/pjms.40.2(ICON).8966)
- [12] R. Rupp and H. J. Gerner, *Neuroprosthetics of the upper extremity — clinical application in spinal cord injury and challenges for the future*. Vienna: Springer Vienna, 2007, pp. 419–426. [Online]. Available: https://doi.org/10.1007/978-3-211-33079-1_55
- [13] W. Young, “Electrical stimulation and motor recovery,” *Cell Transplantation*, vol. 24, no. 3, pp. 429–446, 2015, PMID: 25646771. [Online]. Available: <https://doi.org/10.3727/096368915X686904>
- [14] A. Masdar, B. K. K. Ibrahim, and M. M. Abdul Jamil, “Development of wireless-based low-cost current controlled stimulator for patients with spinal cord injuries,” in *2012 IEEE-EMBS Conference on Biomedical Engineering and Sciences*, 2012, pp. 493–498. [Online]. Available: <https://doi.org/10.1109/IECBES.2012.6498175>
- [15] H.-P. Wang, A.-W. Guo, Y.-X. Zhou, Y. Xia, J. Huang, C.-Y. Xu, Z.-H. Huang, X.-Y. Lü, and Z.-G. Wang, “A wireless wearable surface functional electrical stimulator,” *International Journal of Electronics*, vol. 104, no. 9, pp. 1514–1526, 2017. [Online]. Available: <https://doi.org/10.1080/00207217.2017.1312708>
- [16] T. Keller, M. Lawrence, A. Kuhn, and M. Morari, “New multi-channel transcutaneous electrical stimulation technology for rehabilitation,” in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, 2006, pp. 194–197. [Online]. Available: <https://doi.org/10.1109/IEMBS.2006.259399>
- [17] S. V. C. y Jose Hernandez Jacquez y Mario Cepeda Rubio y Carlos Juárez Rodríguez y Diana Piña Santos y Geshel Guerrero Lopez, “Efecto de los electroestimuladores tens en los síntomas motores de la enfermedad de parkinson,” *Medicina e Investigación Universidad Autónoma del Estado de México*, vol. 11, no. 1, pp. 1–5, 2023. [Online]. Available: <https://doi.org/10.36677/medicinainvestigacion.v11i1.20293>
- [18] F. C. Atamaz, B. Durmaz, M. Baydar, O. Y. Demircioglu, A. Iyiyapici, B. Kuran, S. Oncel, and O. F. Sendur, “Comparison of the efficacy of transcutaneous electrical nerve stimulation, interferential currents, and shortwave diathermy in knee osteoarthritis: A double-blind, randomized, controlled, multicenter study,” *Archives of Physical Medicine and Rehabilitation*, vol. 93, no. 5, pp. 748–756, 2012. [Online]. Available: <https://doi.org/10.1016/j.apmr.2011.11.037>
- [19] S. Farahmand, H. Vahedian, M. Abedinkhan Eslami, and A. M. Sodagar, “Wearable, battery-powered, wireless, programmable

- 8-channel neural stimulator,” in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012, pp. 6120–6123. [Online]. Available: <https://doi.org/10.1109/EMBC.2012.6347390>
- [20] Q. Xu, T. Huang, J. He, Y. Wang, and H. Zhou, “A programmable multi-channel stimulator for array electrodes in transcutaneous electrical stimulation,” in *The 2011 IEEE/ICME International Conference on Complex Medical Engineering*, 2011, pp. 652–656. [Online]. Available: <https://doi.org/10.1109/ICCME.2011.5876821>
- [21] M. Sugandi, H. Satria, H. Arif, N. Nelmiawati, and I. H. Mulyadi, “Low cost wireless ecg patch using esp32,” *Journal Integrasi*, vol. 12, no. 1, pp. 31–35, 2020. [Online]. Available: <https://doi.org/10.30871/ji.v12i1.1764>
- [22] J. Ferdous, B. Roy, M. Hossen, and M. I. Md., “Implementation of iot based patient health monitoring system using esp32 web server,” *International Journal of Advanced Research (IJAR)*, vol. 11, pp. 716–726, 2023. [Online]. Available: <https://dx.doi.org/10.21474/IJAR01/17119>
- [23] K. Sangeethalakshmi, A. S. Preethi, U. Preethi, S. Pavithra, and P. V. Shanmuga, “Patient health monitoring system using iot,” *Materials Today: Proceedings*, vol. 80, pp. 2228–2231, 2023, sI:5 NANO 2021. [Online]. Available: <https://doi.org/10.1016/j.matpr.2021.06.188>
- [24] R. Priyanka and M. Reji, “Iot based health monitoring system using blynk app,” *International Journal of Engineering and Advanced Technology (IJEAT)*, vol. 8, no. 6, pp. 78–81, 2019. [Online]. Available: <http://www.doi.org/10.35940/ijeat.E7467.088619>
- [25] Kodular. (2023) Home – kodular. [Online]. Available: <https://upsalesiana.ec/ing33ar4r25>
- [26] M. D. E. Goodyear, K. Krleza-Jeric, and T. Lemmens, “The declaration of helsinki,” *BMJ*, vol. 335, no. 7621, pp. 624–625, 2007. [Online]. Available: <https://doi.org/10.1136/bmj.39339.610000.BE>
- [27] X. Bao, Y. Zhou, Y. Wang, J. Zhang, X. Lü, and Z. Wang, “Electrode placement on the forearm for selective stimulation of finger extension/flexion,” *PLOS ONE*, vol. 13, no. 1, pp. 1–22, 01 2018. [Online]. Available: <https://doi.org/10.1371/journal.pone.0190936>
- [28] A. Takahashi, J. Brooks, H. Kajimoto, and P. Lopes, “Increasing electrical muscle stimulation’s dexterity by means of back of the hand actuation,” in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI ’21. New York, NY, USA: Association for Computing Machinery, 2021. [Online]. Available: <https://doi.org/10.1145/3411764.3445761>



CONTACT PATCHES OF RADIAL TIRES WITH DIFFERENT LENGTH-TO-WIDTH RATIOS UNDER STATIC LOADS

PARCHES DE CONTACTO DE NEUMÁTICOS RADIALES CON DIFERENTES RELACIONES DE LONGITUD Y ANCHURA BAJO CARGA ESTÁTICA

Fengliang Qiao¹ , Zhaojie Shen^{2,*} , Yuxia Kang³ 

Received: 26-01-2024, Received after review: 09-10-2024, Accepted: 05-11-2024, Published: 01-01-2025

Abstract

The aspect ratio of tires significantly influences the tread contact patch, which is closely related to the vehicle's driving performance and handling. This study investigates the effect of radial tires with varying aspect ratios on contact patches under different loads and inflation pressures. The size and shape of the contact patches, along with the pressure distribution in tires with different aspect ratios, were analyzed. Five finite element tire models with aspect ratios of 55%, 60%, 65%, 70%, and 75% were developed. The simulation models of 205/55R16 were validated against experimental results. The findings reveal that as the aspect ratio increases, the contact length along the tire's axial direction decreases, while the contact width along the rolling direction increases. Minimal differences in contact area were observed among tires with different aspect ratios under the same static load. For a given load, as the length-to-width ratio increases, the tread width of the contact patch decreases, while its length increases. Additionally, with an increasing length-to-width ratio, the contact patch shape transitions from a saddle to a barrel-like form. The maximum normal contact stress occurs at the shoulder of the tire for aspect ratios of 55%, 60%, and 65%, but shifts to the center of the tread for aspect ratios of 70% and 75%. The primary influence of the aspect ratio is on the contact size.

Keywords: bumper, cabuya fiber, fiberglass, optimization, resin, simulation, VARTM

Resumen

La relación de aspecto de los neumáticos influye significativamente en el área de contacto de la banda de rodadura, afectando el rendimiento de conducción y manejo del vehículo. Este estudio analiza el efecto de neumáticos radiales con diferentes relaciones de longitud y diámetro en las manchas de contacto bajo diversas cargas o presiones de inflado. Se evaluaron el tamaño, forma y distribución de presión en neumáticos con relaciones de aspecto de 55 %, 60 %, 65 %, 70 % y 75 %, mediante cinco modelos de elementos finitos. El modelo 205/55R16 fue validado experimentalmente. Los resultados indican que, al aumentar la relación de aspecto, la longitud de contacto a lo largo del eje del neumático disminuye, mientras que la anchura en la dirección de rodadura aumenta. Bajo la misma carga estática, el área de contacto varía poco entre relaciones de aspecto diferentes. Sin embargo, con mayor relación longitud-anchura, el ancho de la banda de rodadura disminuye y la longitud aumenta. Asimismo, la forma de la mancha de contacto cambia de silla de montar a tambor de cintura. Los máximos valores de tensión de contacto normal se localizan en los hombros para relaciones de aspecto de 55 %, 60 % y 65 %, y en el centro de la banda para relaciones de 70 % y 75 %. La principal influencia de la relación de aspecto radica en el tamaño del contacto.

Palabras clave: Relación vertical y horizontal, parches de contacto, tensión de contacto, neumáticos radiales

¹Shanghai First Branch, Hua Yu Automotive Systems Co., LTD., China.

^{2,*}School of Automobile Engineering, Harbin Institute of Technology, Weihai, China.
Corresponding author ✉: shenzj@hitwh.edu.cn.

³Process Technology Department, Guangxi Yuchai Machinery CO., LTD., China.

Suggested citation: Qiao, F., Shen, Z. and Kang, Y. "Contact patches of radial tires with different length-to-width ratios under static loads," *Ingenius, Revista de Ciencia y Tecnología*, N.º 33, pp. 49-59, 2025, DOI: <https://doi.org/10.17163/ings.n33.2025.05>.

1. Introduction

Tires are the only components of a vehicle that directly contact the road. They convert torque into force at the contact patch, enabling vehicle motion and significantly influencing driving and handling performance [1, 2]. The tire's contact patch, which is closely associated with rolling resistance [3, 4], directly impacts energy losses, fuel consumption, greenhouse gas emissions, and tread durability. Consequently, the grounding performance of tires has been extensively studied, with particular focus on the size and shape of the tire-road contact patch [5–9]. The finite element (FE) method and advanced experimental techniques have been widely employed to analyze tire grounding performance [10–13]. Guo et al. improved the UniTire side force model [14], demonstrating that the contact patch plays a critical role in cornering stiffness and lateral relaxation length. Patrick et al. [15] developed a 3D brush model to characterize contact patch size and pressure distribution. Alobaid et al. [16] enhanced an in-plane rigid-elastic-coupled tire model by incorporating the wheel's vertical movement as a rigid body. This model connects road irregularities as input to the contact patch and outputs the wheel's vertical motion. Additionally, Fathi et al. [17] established a tire finite element model using the Arbitrary Lagrangian-Eulerian formulation for steady-state analysis. However, these models often require substantial computational resources and numerous parameter assumptions, presenting challenges for practical applications.

To validate the simulation results, pressure measurement film systems [18–22] have been developed to capture ground pressure distribution at the tire-road interface and analyze the relationships between imprint geometry characteristics and tread wear. Additionally, computer vision techniques and laser tire sensors have been employed to more accurately measure contact patches and tire deflection [23, 24]. Tomaraee et al. [25] utilized a well-equipped single-wheel tester to examine the relationship between imprint size, load, and inflation pressure. Swami et al. [26] developed a Stereo-Digital Image Correlation (DIC) application to generate three-dimensional (3D) visualizations of contact patches and extract dimensional information. Xie et al. [27] established a 3D finite element (FE) tire-pavement model to study the influence of bias-ply and radial tires on contact stress distribution. Their results demonstrated that the maximum contact stress between a radial tire and the road surface exceeds that of a bias-ply tire. Wang et al. [28] analyzed the effect of friction coefficient on horizontal contact stress using an improved simulation model that incorporates a neo-Hookean rubber material and a multilayer pavement structure. Oubahdou et al. [29] examined tire-pavement contact by employing a realistic description of normal sections and bends, providing detailed analyses of con-

tact patch shape, stress distribution, and surface shear stresses at the tire-pavement interface. Jaime et al. [30] investigated the influence of pavement stiffness on 3D contact stress using a validated FE model. Their findings indicated that while pavement stiffness minimally affects normal contact stress, it significantly impacts longitudinal contact stress. Using a nonlinear FE tire model, Liu et al. [31] developed a tool for fast and accurate prediction of non-uniform tire-pavement contact stresses, integrating deep learning techniques. Relationships between wheel loading and contact patch behavior have been widely explored [32–34]. Zhang et al. [35] constructed a contact model for patterned tire-asphalt pavement, analyzing contact behavior under static load and ABS conditions. Their results revealed that the contact area during braking is 7.7% smaller than that under static loading. Comparisons of contact pressure distributions under static and rolling conditions using FE tire models have shown that braking, acceleration, and cornering significantly impact contact stress distributions, with contact patch length increasing as tire speed decreases [36, 37]. Gu et al. [38] concluded that the longitudinal contact patch length decreases with higher inflation pressure. The contact patch width must be considered for lateral stress analysis [39].

The contact patch plays a crucial role in tire performance, influencing tread wear, steering responsiveness, wet and dry handling, traction, tire noise, and ride quality. Therefore, investigating the tire contact patch is essential for optimizing tire performance. Among the factors affecting the contact patch, the tire aspect ratio is a significant contributor. However, limited studies have explored the relationship between the tire aspect ratio and the contact patch. This study aims to examine the effects of tire aspect ratio on the contact patch under various vertical load and inflation pressure conditions. Finite element (FE) models of the 205/55R16 tire were developed using ABAQUS 6.14 and validated by comparing simulation results of load and deflection with experimental data. Additional FE tire models with aspect ratios of 55%, 60%, 65%, 70%, and 75% were also established. The study analyzed the imprint shapes and contact stress distributions across different vertical loads and inflation pressures.

The article is structured as follows: it begins with a description of the tire materials and modeling process, followed by the validation of the FE models. The discussion section then explores the relationship between tire aspect ratio and the contact patch under static conditions.

2. Materials and methods

The finite element (FE) model and cross-section of the 205/55R16 tire are illustrated in Figure 1. The

width of the tire's cross-section is 205 mm, with an aspect ratio (height-to-width) of 55%. The rim diameter measures 406.4 mm, and the wheel radius is 315.95 mm. Additional FE models with aspect ratios of 60%, 65%, 70%, and 75% were developed by adjusting the cross-section width while maintaining the tire height constant.

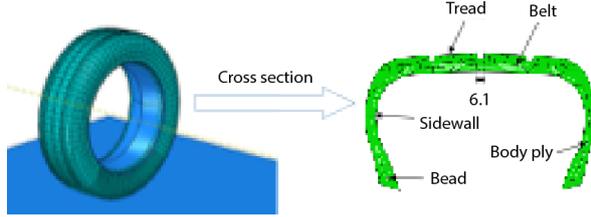


Figure 1. Simulation model of tire 205/55R16

The pavement and rim were simplified as rigid surfaces since their stiffness is significantly higher than that of the tire. To ensure higher mesh quality, only the longitudinal tread pattern was considered in the tire model, while the transverse tread pattern was omitted. The two-dimensional cross-sections of the tire were created using AutoCAD software and subsequently imported into ABAQUS to develop the 3D models. Two belts were incorporated into the simulation: Belt 1, located closer to the tread, had a width of 94 mm, while Belt 2 had a width of 114 mm. A concentrated force along the Z-axis was applied at the center point of the rim. The discretization method for tire-pavement interaction was utilized to predict surface-to-surface contact. A penalty function was employed to simulate tangential contact behavior, with a friction coefficient set at 0.8. Additionally, the grounding mesh was refined to enhance the accuracy of the modeling predictions.

The hyperelastic mechanical properties of rubber materials are described using the Mooney-Rivlin constitutive model [40]. The corresponding material properties are provided in Table 1. In this study, rubber deformation is modeled as uniform deformation of isotropic hyperelastic bodies. Consequently, the strain energy density is expressed using strain invariants as follows:

Table 1. Mooney-Rivlin material properties

Components	C_{10}	C_{01}	Density (kg/m^3)
Tread	0.5792	0.1448	1112
Sidewall	0.5240	0.1310	1110.1
Belt	10,848	0.2712	1144
Body ply	0.6159	0.1540	1058
Bead	11500	2870	10007

Where $I_3 = 1$ Given that the material is fully incompressible, equation (1) simplifies to:

$$W(E) = W(I_1, I_2, I_3) \quad (1)$$

Where, C_{ij} represent the material coefficients. For practical engineering applications, equation (2) is further reduced to:

$$W_R = \sum_{i,j=0}^{\infty} C_{ij}(I_1 - 3)^i (I_2 - 3)^j \quad (2)$$

$$W_{MR} = C_{10}(I_1 - 3) + C_{01}(I_2 - 3) \quad (3)$$

The rubber-cord composite was simulated using rebar elements to account for both rubber and cord materials simultaneously, with each material meshed separately. The properties of the rebar material are listed in Table 2. Inflation pressure was applied to the inner surface of the tire to simulate various inflation conditions. According to the GB/T 2977-2016 standard, the parameters of the tire are defined as:

Where, B is the designed tire section width, B^* is the inflated tire section width, H is the section height, D is the rim diameter, D^* is the inflated tire diameter.

$$\begin{cases} B \times 1.05 \geq B^* \geq B \times 0.95 \\ 2 \times H \times 1.03 + D \geq D^* \geq 2 \times H \times 0.97 + D \end{cases} \quad (4)$$

Table 2. Rebar material properties

Rebar	Elastic modulus (MPa)	Poisson ratio	Density (kg/m^3)	Orientation (deg)
Belt 1	110530	0.4	7804	66.25
Belt 2	110530	0.4	7804	-66.25
Body cord	2710	0.4	1251	0

2.1. Model validation

To validate the model, the relationship between tire deflection and static load was compared with experimental data from reference [41]. The experiments were conducted using a 5-in-1 Tire Stiffness Testing Machine (ke · TEK, Taiwan Hung Ta Instrument Co., Ltd) with a maximum displacement of 1500 mm and an accuracy of ± 0.1 mm. The tested tire was a 205/55 R16, with a tire pressure of 0.24 MPa. Radial loads of 3000 N, 3500 N, 4000 N, and 5000 N were applied. The same parameters were used in the finite element (FE) simulations. The largest error between the simulation and experimental results was 6.8%, indicating that the simplified FE tire model is reasonable and suitable for investigating the relationship between aspect ratio and tire-ground contact patch. Simulation and experimental results for deflection under different static loads are presented in Table 3.

Table 3. Simulation and experimental results with relative error

Load (N)	Deflection		Relative error (%)
	Simulation (mm)	Experiment (mm)	
3000N	15.1	15	0.7
3500N	16.9	17.5	3.4
4000N	18.6	19.8	6.1
5000N	21.9	23.5	6.8

3. Results and Discussion

3.1. The size and shape of the contact patch under varying vertical loads

Figure 2 illustrates the contact patches of tires with varying aspect ratios under vertical loads of 3000 N, 4000 N, and 5000 N.

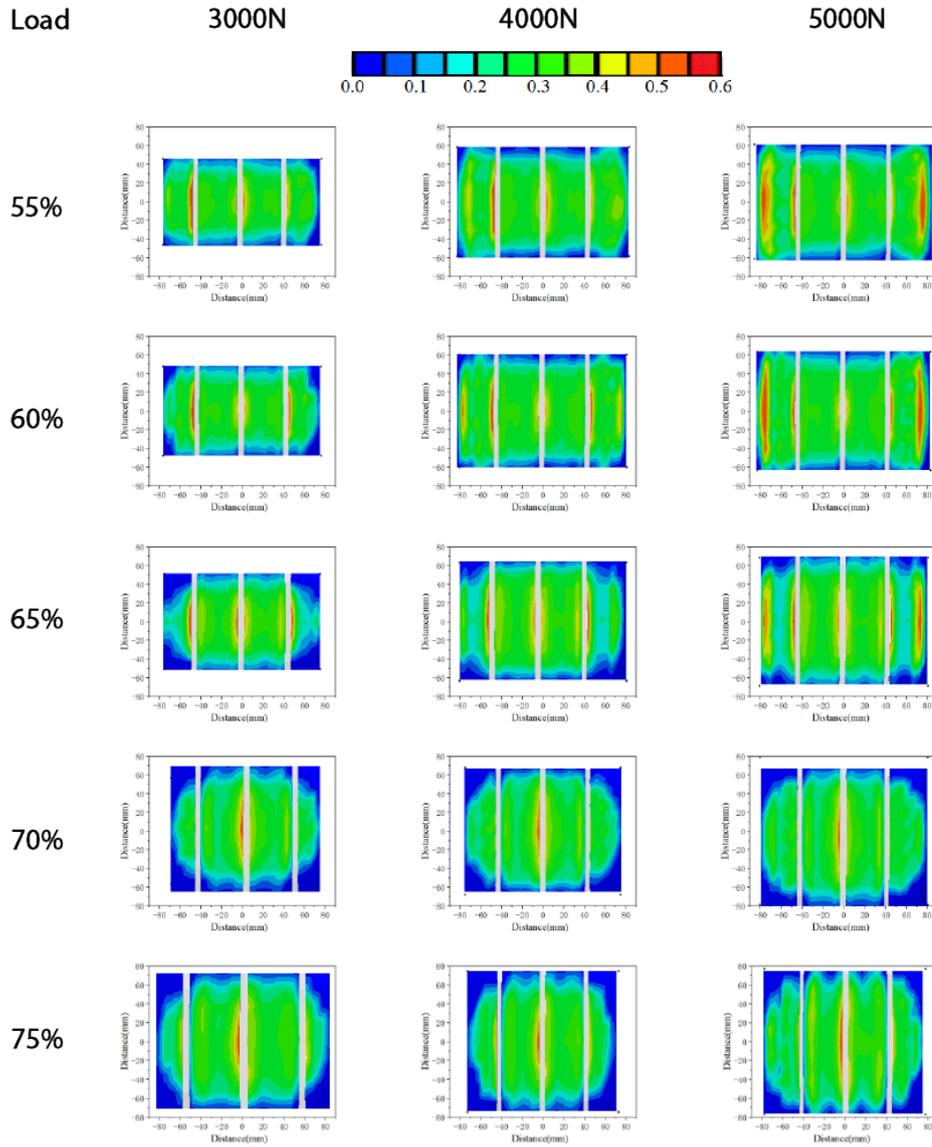


Figure 2. The shape of the contact patches of tires with different aspect ratios under varying vertical loads

The inflation pressure was set to 0.24 MPa. Figure 2 shows that the contact patch exhibited a saddle shape for aspect ratios of 55% and 60%, transitioned to a rectangular shape for an aspect ratio of 65%, and evolved into a waist-drum shape for aspect ratios of 70% and 75%. To better characterize the contact

patch, the tread contact length and width were measured. The tread contact length refers to the maximum length of the contact patch along the tire's rolling direction, while the tread contact width represents the maximum width of the contact patch along the tire's axial direction.

The contact areas for different aspect ratios, obtained from ABAQUS simulations, are presented in Table 4. The results indicate that as the aspect ratio increases, the tread contact width decreases, while the tread contact length increases under the same static load. Additionally, for a given aspect ratio, both the tread contact width and length, as well as the overall contact patch area, increase with higher vertical loads.

Table 4. The size of the contact patch of tires with different aspect ratios under different vertical loads

Load (N)	Aspect ratio (%)	Width (mm)	Length (mm)	Area of contact Patch (cm ²)
3000	55	153.5	92.2	111.5
	60	153.0	95.4	114.2
	65	150.4	101.7	112.1
	70	137.3	113.9	118.5
	75	128.4	123.1	107.1
4000	55	166.6	116.9	150.7
	60	161.7	120.7	147.6
	65	161.2	126.8	153.3
	70	150.0	135.7	152.8
	75	145.9	149.0	151.5
5000	55	171.2	122.8	166.7
	60	166.1	127.0	171.9
	65	161.3	137.2	174.8
	70	160.6	157.8	186.6
	75	155.5	153.9	187.5

Figure 3 illustrates the relationships between the aspect ratio and the contact patch area under varying vertical loads. The results indicate that the contact patch area increases as the vertical load increases. Under loads of 3000 N and 4000 N, the contact patch area fluctuates with increasing aspect ratio, whereas under a load of 5000 N, it consistently increases with higher aspect ratios. The relationship between the tire aspect ratio and the contact patch area is non-linear. A larger contact patch area generally results in reduced vibration and noise, as well as improved grip performance. However, an increased contact patch area also leads to higher average tread wear and greater energy loss. Across all three vertical loads, the tire with an aspect ratio of 55% exhibited a relatively smaller contact area.

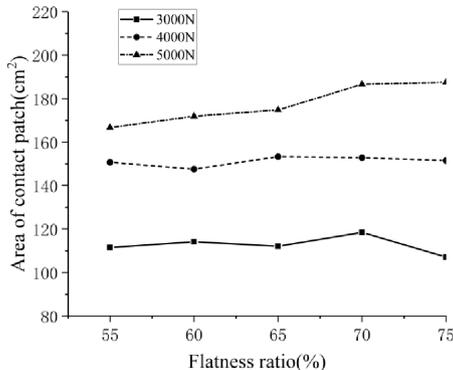


Figure 3. Relation curves between the aspect ratio and the contact patch area under different loads

3.2. Contact stress distribution under different vertical loads

Figure 4 presents the contact stress distribution of tires along the node path for different aspect ratios under varying vertical loads.

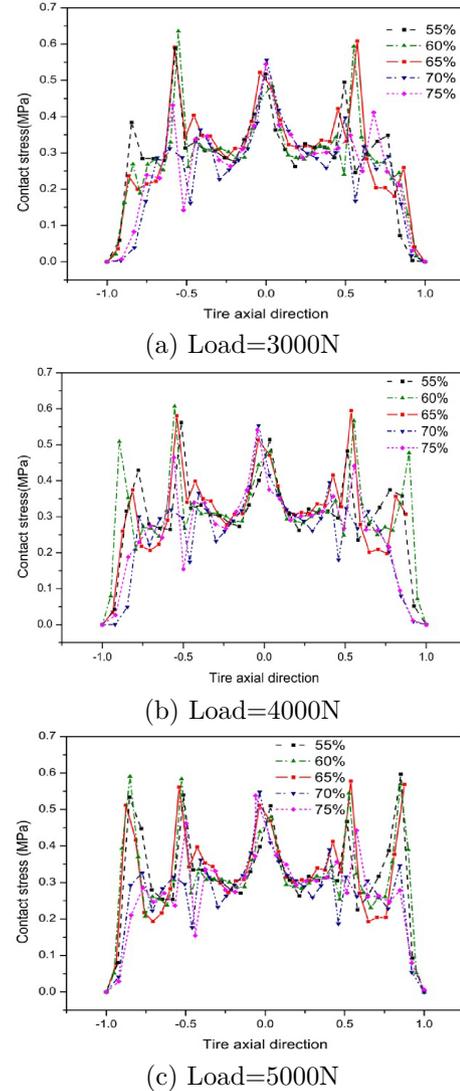


Figure 4. Contact stress distribution of tires with different aspect ratios under different loads

The contact stress curves along the tread width exhibit a multi-peak pattern, with peak values primarily concentrated at the shoulder and longitudinal groove. This pattern arises from stress concentrations caused by abrupt changes in tread geometry at the shoulder and longitudinal groove. It was due to stress concentration caused by dramatic changes in tread shape at the shoulder and groove. For the vertical load of 3000N as shown in Figure 4a, the peak value of contact stress was mainly distributed in the tread groove, and the peak values of tire with aspect ratio of 60% and 65% were larger than that of other aspect ratio tires. For

the vertical load of 4000N as shown in Figure 4b, the contact stress increased obviously at the shoulder. The maximum contact stress was at the shoulder of the tire with aspect ratio of 60%.

For a vertical load of 5000N, as shown in Figure 4c, the contact stress on both sides of the tread decreased with an increase in the aspect ratio, while the contact stress at the center of the tread increased. The maximum contact stress was observed at the shoulder of the tire for aspect ratios of 55%, 60%, and 65%. In contrast, for aspect ratios of 70% and 75%, the maximum stress shifted to the central part of the tread. Regions with higher contact stress typically experience more severe wear [42]. Additionally, as the center of the tread is thinner, aspect ratios of 70% and 75% may reduce the tire's service life.

3.3. Size and shape of the contact patch under different inflation pressures

The sizes of the contact patches for tires under varying inflation pressures are presented in Table 5. The width of the contact patch decreased progressively as the aspect ratio increased, while the length of the contact patch increased under the same inflation pressure. Both the width and length of the contact patch, as well as its overall area, decreased with increasing inflation pressure for a given aspect ratio. Figure 5 illustrates the contact patches of tires with different aspect ratios under inflation pressures of 0.20, 0.24, and 0.28 MPa. The vertical load was set at 3000 N. The results show that the shape of the contact patch transitioned from a saddle shape to a waist drum shape as the aspect ratio increased under the same inflation pressure. Additionally, as the inflation pressure increased, the distribution of contact stress became more uneven.

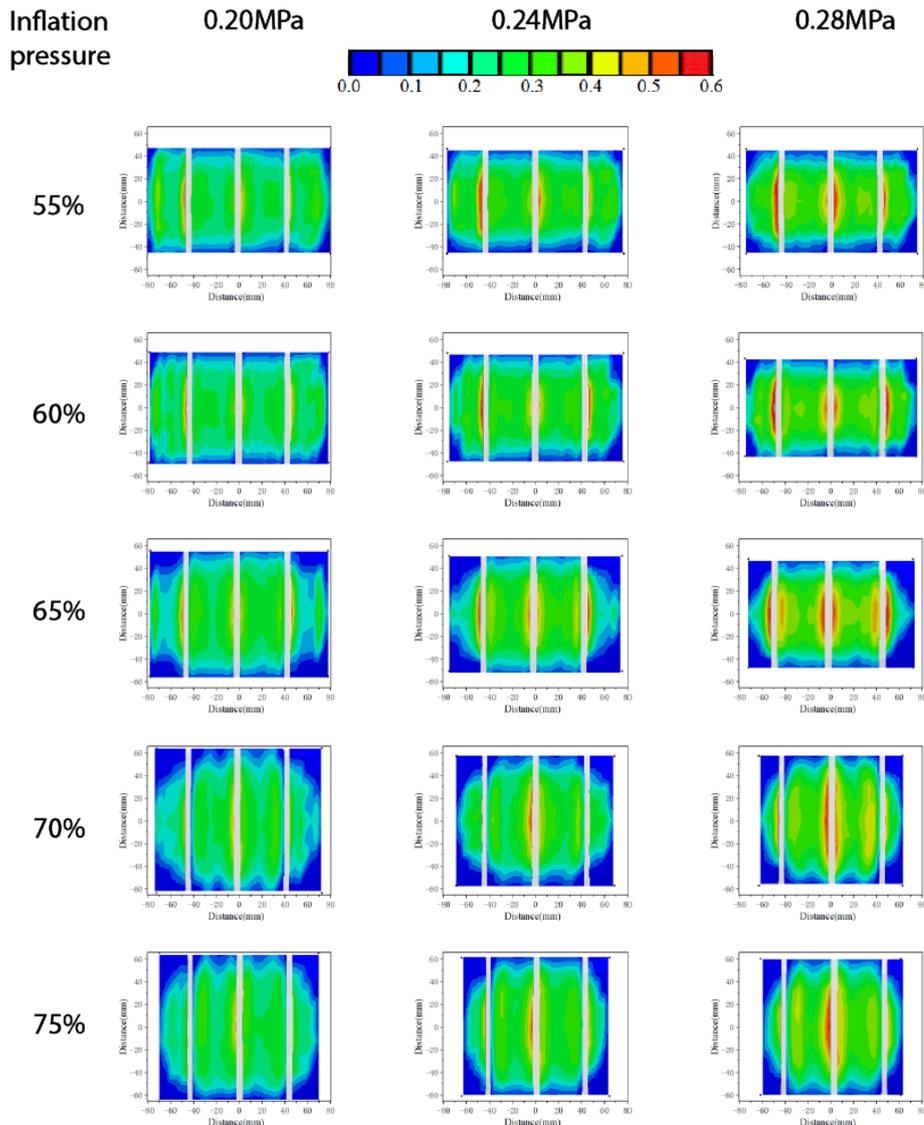


Figure 5. Contact patch shapes of tires with varying aspect ratios under different inflation pressures

Table 5. Size of the contact patch for tires with different aspect ratios under various inflation pressures

Inflation pressure (MPa)	Aspect ratio (%)	Width (mm)	Length (mm)	Area of contact patch (cm ²)
0.20	55	160.8	93.9	124.0
	60	157.1	97.2	122.8
	65	155.7	110.9	124.6
	70	144.9	127.8	135.8
	75	139.4	131.2	135.7
0.24	55	153.5	92.4	111.5
	60	153.0	95.4	114.2
	65	150.4	101.7	112.1
	70	137.3	113.9	118.5
	75	128.4	123.1	107.1
0.28	55	150.1	92.2	106.7
	60	144.1	95.6	105.0
	65	130.9	109.2	100.6
	70	125.8	114.2	99.4
	75	123.0	120	100.2

Figure 6 illustrates the relationship between the aspect ratio and the contact patch area under various inflation pressures. The data reveal that the contact patch area decreases as inflation pressure increases. At an inflation pressure of 0.2 MPa, the contact patch area increases as the aspect ratio rises from 60% to 70%. For an inflation pressure of 0.24 MPa, the contact patch area fluctuates with increasing aspect ratio. Conversely, at an inflation pressure of 0.28 MPa, the contact patch area decreases as the aspect ratio increases. Across all three inflation pressure conditions, tires with an aspect ratio of 55% consistently exhibited relatively smaller contact patch areas.

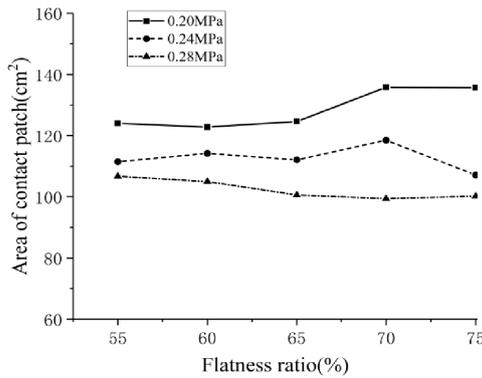
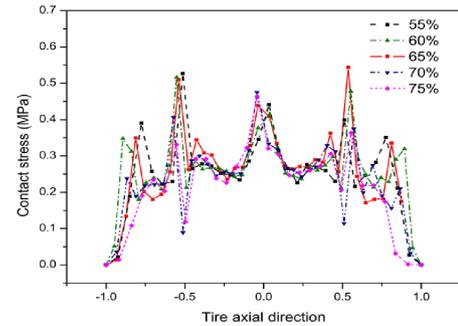


Figure 6. Relation curves between aspect ratio and contact patch area under different inflation pressures

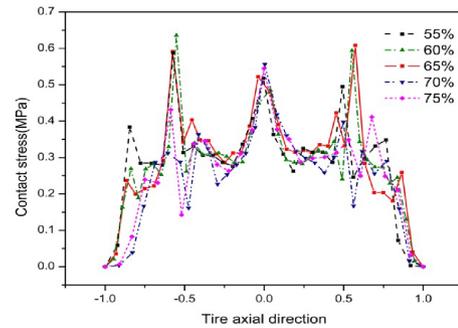
3.4. Contact stress distribution under different inflation pressures

Contact pressure is a critical factor in vehicle performance, directly influencing grip and traction. These parameters are closely associated with a vehicle’s acceleration, braking, and cornering capabilities. While higher contact pressure generally enhances grip, excessive pressure can accelerate tire wear prematurely.

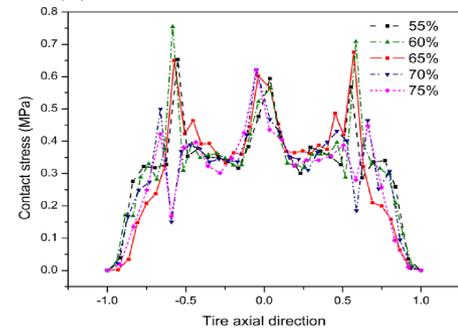
Figure 7 illustrates the contact stress distribution of tires along the node path for different aspect ratios under varying inflation pressures. The contact stress curves along the tread width exhibited a multi-peak pattern. As inflation pressure increased, contact stress at the shoulders decreased, while stress at the longitudinal grooves increased. When inflation pressure was set to 0.20 MPa (Figure 7a), the peak contact stress remained below 0.6 MPa. In contrast, at inflation pressures of 0.24 MPa and 0.28 MPa (Figures 7b and 7c), the peak contact stress exceeded 0.6 MPa. Across all inflation pressures, the peak contact stress at the grooves on both sides of the tread for tires with aspect ratios of 60% and 65% was higher than those of other aspect ratios. Conversely, the contact stress at the middle groove of the tread for aspect ratios of 60% and 65% was lower than those observed in other aspect ratios under identical inflation conditions.



(a) Inflation pressure=0.20MPa



(b) Inflation pressure=0.24MPa



(c) Inflation pressure=0.28MPa

Figure 7. Contact stress distribution of tires with different aspect ratios under varying inflation pressures

Uneven contact pressure distribution leads to irregular tire wear. As shown in Figure 7a, tires with aspect ratios of 70% and 75% exhibit lower and more uniform contact stress compared to those with aspect ratios of 55%, 60%, and 65%. This suggests a higher likelihood of uneven wear for tires with lower aspect ratios. When inflation pressure increased from 0.2 MPa to 0.24 MPa (Figure 7b), this disparity became more pronounced. Additionally, the contact stress at the shoulder of tires with lower aspect ratios is higher than that of tires with higher aspect ratios, indicating better stability and handling during cornering. The contact stress at the center of the tread is nearly identical across all aspect ratios, with the primary differences occurring at the shoulder and longitudinal grooves.

4. Conclusions

This study presents finite element (FE) models of 205/XX R16 tires developed using ABAQUS software to investigate the size and shape of the contact patch, as well as the axial distribution of contact pressure within the contact patch for tires with different aspect ratios under varying static wheel loads and inflation pressures. The main conclusions are as follows:

1. The FE simulation results for tire deflection under different static loads align closely with experimental data, with a maximum error of 6.8%, confirming the accuracy and reliability of the model.
2. The tread width of the contact patch decreased, and the length increased as the aspect ratio increased. The contact patch area fluctuated with increasing aspect ratio under static loads of 3000 N and 4000 N, while it increased consistently under a static load of 5000 N. The shape of the contact patch evolved from a saddle shape to a waist-drum shape as the aspect ratio increased.
3. Lower inflation pressures resulted in larger contact patch areas and a more uniform distribution of contact stress. As inflation pressure increased under the same static wheel load, the contact stress at the center of the tread increased.
4. The contact stress distribution along the axial direction of the tire exhibited a multi-peak pattern, with peaks primarily located at the shoulder and longitudinal grooves. For tires with aspect ratios of 55%, 60%, and 65%, the maximum contact stress was observed at the shoulder. In contrast, for tires with aspect ratios of 70% and 75%, the maximum stress occurred at the center of the tread.

References

- [1] M. Abe, *Vehicle Handling Dynamics (Second Edition)*, second edition ed., M. Abe, Ed. Butterworth-Heinemann, 2015. [Online]. Available: <https://doi.org/10.1016/B978-0-08-100390-9.01001-0>
- [2] Y. Suo, W. Yang, D. Lu, Y. Zhang, and M. Che, "Analysis of camber-caused asymmetric characteristics using finite element method and pure camber semi-empirical modeling," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, p. 09544070241272802, 2024. [Online]. Available: <https://doi.org/10.1177/09544070241272802>
- [3] F. B. Luigi Romano and B. Jacobson, "An extended lugre-brush tyre model for large camber angles and turning speeds," *Vehicle System Dynamics*, vol. 61, no. 6, pp. 1674–1706, 2023. [Online]. Available: <https://doi.org/10.1080/00423114.2022.2086887>
- [4] Z. Shi, Y. M. Mohammed, N. Uddin, and G. Chen, "A vehicle-bridge interaction model considering contact patch size and vehicle self-generated excitation – a theoretical study," *Engineering Structures*, vol. 298, p. 117079, 2024. [Online]. Available: <https://doi.org/10.1016/j.engstruct.2023.117079>
- [5] C. Suvanjumrat and J. Phromjan, "The contact patch characterization of various solid tire testing methods by finite element analysis and experiment," *International Journal of Geomate*, vol. 19, no. 76, pp. 25–32, 2020. [Online]. Available: <https://doi.org/10.21660/2020.76.9134>
- [6] N. Ryzí, R. Stořek, J. Maloch, and M. Střaniřka, "How does heat generation affect the cut and chip wear of rubber?" *Polymer Bulletin*, vol. 81, no. 18, pp. 17 213–17 232, Dec 2024. [Online]. Available: <https://doi.org/10.1007/s00289-024-05498-1>
- [7] Y. Nakajima and S. Hidano, "Theoretical tire model considering two-dimensional contact patch for force and moment," *Tire Science and Technology*, vol. 50, no. 1, pp. 27–60, 07 2021. [Online]. Available: <https://doi.org/10.2346/tire.21.20005>
- [8] J. Prakash, M. Vignati, and E. Sabbioni, "An exponential decay model for decaying of contact patch friction steering moment with rolling speed," *Tire Science and Technology*, vol. 52, no. 1, pp. 34–50, 03 2024. [Online]. Available: <https://doi.org/10.2346/tire.23.21017>

- [9] J. Bastiaan, A. Chawan, W. Eum, K. Alipour, F. Rouhollahi, M. Behroozi, and J. Baqersad, "Intelligent tire prototype in longitudinal slip operating conditions," *Sensors*, vol. 24, no. 9, 2024. [Online]. Available: <https://doi.org/10.3390/s24092681>
- [10] J. M. Conradie, P. S. Els, and P. S. Heyns, "Finite element modelling of off-road tyres for radial tyre model parameterization," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 230, no. 4, pp. 564–578, 2016. [Online]. Available: <https://doi.org/10.1177/0954407015590018>
- [11] M. Zhang, H.-J. Unrau, M. Gießler, and F. Gauterin, "A detailed tire tread friction model considering dynamic friction states," *Tribology International*, vol. 193, p. 109342, 2024. [Online]. Available: <https://doi.org/10.1016/j.triboint.2024.109342>
- [12] Z. Liu, F. Wang, Z. Cai, Y. Wei, and S. Marburg, "A novel theoretical model of tire in-plane dynamics on uneven roads and its experimental validation," *Mechanical Systems and Signal Processing*, vol. 186, p. 109854, 2023. [Online]. Available: <https://doi.org/10.1016/j.ymsp.2022.109854>
- [13] P. Millan and J. Ambrósio, "Tire–road contact modelling for multibody simulations with regularised road and enhanced ua tire models," *Multibody System Dynamics*, Apr 2024. [Online]. Available: <https://doi.org/10.1007/s11044-024-09987-z>
- [14] Guo, Konghui, Chen, Ping, Xu, Nan, Yang, Chao, and Li, Fei, "Tire side force characteristics with the coupling effect of vertical load and inflation pressure," *SAE International Journal of Vehicle Dynamics, Stability, and NVH*, vol. 3, no. 1, pp. 19–30, nov 2018. [Online]. Available: <https://doi.org/10.4271/10-03-01-0002>
- [15] P. Riehm, H.-J. Unrau, F. Gauterin, S. Torbrügge, and B. Wies, "3d brush model to predict longitudinal tyre characteristics," *Vehicle System Dynamics*, vol. 57, no. 1, pp. 17–43, 2019. [Online]. Available: <https://doi.org/10.1080/00423114.2018.1447135>
- [16] F. Alobaid and S. Taheri, "The modified in-plane rigid-elastic-coupled tire modal model: dynamic response to short wavelength road profiles," *Vehicle System Dynamics*, vol. 62, no. 12, pp. 3076–3097, 2024. [Online]. Available: <https://doi.org/10.1080/00423114.2024.2316683>
- [17] H. Fathi, Z. El-Sayegh, and M. H. R. Ghorishy, "Prediction of rolling resistance and wheel force for a passenger car tire: A comparative study on the use of different material models and numerical approaches," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, p. 09544070241244556, 2024. [Online]. Available: <https://doi.org/10.1177/09544070241244556>
- [18] X. Gao, Y. Wang, W. Fan, Z. Long, X. Li, X. Yue, Y. Liu, Y. Yan, and J. Wang, "Modeling and experimental verification of torsional deformation constitutive model of tread rubber based on digital image correlation," *Experimental Techniques*, vol. 47, no. 4, pp. 749–765, Aug 2023. [Online]. Available: <https://doi.org/10.1007/s40799-022-00583-4>
- [19] S. K. Pradhan, A. S. Rathore, S. Sehgal, P. Sonia, G. Ramu, and C. Prakash, "Development and validation of test rig for experimental analysis of contact behavior between rail wheel-rail and rubber tire-rail in road cum rail vehicles," *Indian Journal of Engineering and Materials Sciences (IJEMS)*, vol. 31, no. 1, pp. 84–92, 2024. [Online]. Available: <https://doi.org/10.56042/ijems.v31i1.986>
- [20] J. Guan, X. Zhou, L. Liu, M. Ran, and Y. Yan, "Investigation of tri-axial stress sensing and measuring technology for tire-pavement contact surface," *Coatings*, vol. 12, no. 4, 2022. [Online]. Available: <https://doi.org/10.3390/coatings12040491>
- [21] T. Saisaengtham, J. Phromjan, R. Rugsaj, S. Phakdee, and C. Suvanjumrat, "Pavement-tire contact patch effects on air volume using finite element method," *International Journal of Geomate*, vol. 26, no. 113, pp. 50–57, 2024. [Online]. Available: <https://doi.org/10.21660/2024.113.g13179>
- [22] C. Liang, D. Zhu, G. Wang, and M. Shan, "Experimental study on tire-road dynamic contact pressure distribution using ftir imaging," *International Journal of Automotive Technology*, vol. 22, no. 5, pp. 1305–1317, Oct 2021. [Online]. Available: <https://doi.org/10.1007/s12239-021-0114-3>
- [23] N. Xu, H. Askari, Y. Huang, J. Zhou, and A. Khajepour, "Tire force estimation in intelligent tires using machine learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 4, pp. 3565–3574, 2022. [Online]. Available: <https://doi.org/10.1109/TITS.2020.3038155>
- [24] N. Xu, J. Zhou, B. H. G. Barbosa, H. Askari, and A. Khajepour, "A soft sensor for estimating tire cornering properties for intelligent tires," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 53, no. 10, pp. 6056–6066, 2023. [Online]. Available: <https://doi.org/10.1109/TSMC.2023.3281474>

- [25] P. Tomaraee, A. Mardani, A. Mohebbi, and H. Taghavifar, "Relationships among the contact patch length and width, the tire deflection and the rolling resistance of a free-running wheel in a soil bin facility," *Spanish Journal of Agricultural Research*, vol. 13, no. 2, p. e0211, May 2015. [Online]. Available: <https://doi.org/10.5424/sjar/2015132-5245>
- [26] A. Swami, C. Liu, J. Kubenz, G. Prokop, and A. K. Pandey, "Experimental study on tire contact patch characteristics for vehicle handling with enhanced optical measuring system," *SAE International Journal of Vehicle Dynamics Stability and NVH*, vol. 5, no. 3, pp. 333–350, 2021. [Online]. Available: <https://doi.org/10.4271/10-05-03-0023>
- [27] Y. Xie and Q. Yang, "Tyre-pavement contact stress distribution considering tyre types," *Road Materials and Pavement Design*, vol. 20, no. 8, pp. 1899–1911, 2019. [Online]. Available: <https://doi.org/10.1080/14680629.2018.1473285>
- [28] Y. Wang, Y. Lu, and C. Si, "Tire-pavement coupling dynamic simulation under tire high-speed-rolling condition," *International Journal of Simulation Modelling*, vol. 15, pp. 236–248, 06 2016. [Online]. Available: [http://dx.doi.org/10.2507/IJSIMM15\(2\)4.332](http://dx.doi.org/10.2507/IJSIMM15(2)4.332)
- [29] Y. Oubahdou, E.-R. Wallace, P. Reynaud, B. Picoux, J. Dopeux, C. Petit, and D. NÁ©lias, "Effect of the tire â€“ pavement contact at the surface layer when the tire is tilted in bend," *Construction and Building Materials*, vol. 305, p. 124765, 2021. [Online]. Available: <https://doi.org/10.1016/j.conbuildmat.2021.124765>
- [30] J. A. Hernandez and I. L. Al-Qadi, "Tire-pavement interaction modelling: hyper-elastic tire and elastic pavement," *Road Materials and Pavement Design*, vol. 18, no. 5, pp. 1067–1083, 2017. [Online]. Available: <https://doi.org/10.1080/14680629.2016.1206485>
- [31] X. Liu and I. L. Al-Qadi, "Three-dimensional tire-pavement contact stresses prediction by deep learning approach," *International Journal of Pavement Engineering*, vol. 23, no. 14, pp. 4991–5002, 2022. [Online]. Available: <https://doi.org/10.1080/10298436.2021.1990288>
- [32] J. Ye, Z. Zhang, J. Jin, R. Su, and B. Huang, "Estimation of tire-road friction coefficient with adaptive tire stiffness based on r-sckf," *Nonlinear Dynamics*, vol. 112, no. 2, pp. 945–960, Jan 2024. [Online]. Available: <https://doi.org/10.1007/s11071-023-09088-0>
- [33] Sun, Lihong, Lu, Dang, and Li, Bing, "Analysis and prediction of tire traction properties for different inflation pressures based on vertical deflection control method," *SAE International Journal of Vehicle Dynamics, Stability, and NVH*, vol. 5, no. 3, pp. 307–315, apr 2021. [Online]. Available: <https://doi.org/10.4271/10-05-03-0021>
- [34] Z. Gong, Y. Miao, W. Li, W. Yu, and L. Wang, "Analysis of tyre-pavement contact behaviour of heavy truck in full-scale test," *International Journal of Pavement Engineering*, vol. 24, no. 1, p. 2235630, 2023. [Online]. Available: <https://doi.org/10.1080/10298436.2023.2235630>
- [35] B. Zheng, J. Chen, R. Zhao, J. Tang, R. Tian, S. Zhu, and X. Huang, "Analysis of contact behaviour on patterned tire-asphalt pavement with 3-d fem contact model," *International Journal of Pavement Engineering*, vol. 23, no. 2, pp. 171–186, 2022. [Online]. Available: <https://doi.org/10.1080/10298436.2020.1736294>
- [36] S. T. Pooya Behroozinia and R. Mirzaeifar, "An investigation of intelligent tires using multiscale modeling of cord-rubber composites," *Mechanics Based Design of Structures and Machines*, vol. 46, no. 2, pp. 168–183, 2018. [Online]. Available: <https://doi.org/10.1080/15397734.2017.1321488>
- [37] I. L. A.-Q. Hao Wang and I. Stanciulescu, "Simulation of tyre-pavement interaction for predicting contact stresses at static and various rolling conditions," *International Journal of Pavement Engineering*, vol. 13, no. 4, pp. 310–321, 2012. [Online]. Available: <https://doi.org/10.1080/10298436.2011.565767>
- [38] T. Gu, B. Li, Z. Quan, S. Bei, G. Yin, J. Guo, X. Zhou, and X. Han, "The vertical force estimation algorithm based on smart tire technology," *World Electric Vehicle Journal*, vol. 13, no. 6, 2022. [Online]. Available: <https://doi.org/10.3390/wevj13060104>
- [39] P. Rosca, M. L. Marmureanu, T. V. Tigianescu, C. M. Pîrvulescu, I. M. Bîndac, and C. Doru, "Determination of tyre-ground interaction parameters through image processing in matlab," *International Journal of Heavy Vehicle Systems*, vol. 28, no. 5, pp. 630–649, 2021. [Online]. Available: <https://doi.org/10.1504/IJHVS.2021.120913>
- [40] H. B. Huang, X. D. Yu, J. P. Liu, and Z. Yao, "Asymmetry investigation on radial tire contact pressure distribution," *Chinese Journal of System Simulation*, vol. 30, no. 8, pp. 2991–2998, 2018. [Online]. Available: <https://doi.org/10.16182/j.issn1004731x.joss.201808021>

-
- [41] T. Doi and K. Ikeda, "Effect of tire tread pattern on groove wander of motorcycles," *Tire Science and Technology*, vol. 13, no. 3, pp. 147–153, 07 1985. [Online]. Available: <https://doi.org/10.2346/1.2150992>
- [42] C. Wang, H. Huang, X. Chen, and J. Liu, "The influence of the contact features on the tyre wear in steady-state conditions," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 231, no. 10, pp. 1326–1339, 2017. [Online]. Available: <https://doi.org/10.1177/0954407016671462>



A COMPREHENSIVE EVALUATION OF AI TECHNIQUES FOR AIR QUALITY INDEX PREDICTION: RNNs AND TRANSFORMERS

UNA EVALUACIÓN INTEGRAL DE LAS TÉCNICAS DE IA PARA PREDECIR EL ÍNDICE DE CALIDAD DEL AIRE: RNN Y TRANSFORMERS

Pablo Andrés Buestán Andrade¹ , Pedro Esteban Carrión Zamora¹ ,
 Anthony Eduardo Chamba Lara¹ , Juan Pablo Pazmiño Piedra^{1,*} 

Received: 18-10-2024, Received after review: 15-11-2024, Accepted: 25-11-2024, Published: 01-01-2025

Abstract

This study evaluates the effectiveness of Recurrent Neural Networks (RNNs) and Transformer-based models in predicting the Air Quality Index (AQI). Accurate AQI prediction is critical for mitigating the significant health impacts of air pollution and plays a vital role in public health protection and environmental management. The research compares traditional RNN models, including Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks, with advanced Transformer architectures. Data were collected from a weather station in Cuenca, Ecuador, focusing on key pollutants such as CO , NO_2 , O_3 , $PM_{2.5}$, and SO_2 . Model performance was assessed using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the Coefficient of Determination (R^2). The findings reveal that the LSTM model achieved superior performance, with an R^2 of 0.701, an RMSE of 0.087, and an MAE of 0.056, demonstrating superior capability in capturing temporal dependencies within complex datasets. Conversely, while Transformer-based models exhibited potential, they were less effective in handling intricate time-series data, resulting in comparatively lower accuracy. These results position the LSTM model as the most reliable approach for AQI prediction, offering an optimal balance between predictive accuracy and computational efficiency. This research contributes to improving AQI forecasting and underscores the importance of timely interventions to mitigate the harmful effects of air pollution.

Keywords: Air Quality Index, RNN, LSTM, Transformers, Pollution Forecasting

Resumen

Este estudio evalúa la eficacia de las redes neuronales recurrentes (RNN) y los modelos basados en transformadores para predecir el índice de calidad del aire (ICA). La investigación compara los modelos RNN tradicionales, incluidos los de memoria a corto y largo plazo (LSTM) y la unidad recurrente controlada (GRU), con arquitecturas avanzadas de transformadores. El estudio utiliza datos de una estación meteorológica en Cuenca, Ecuador, centrándose en contaminantes como CO , NO_2 , O_3 , $PM_{2.5}$ y SO_2 . Para evaluar el rendimiento de los modelos, se utilizaron métricas clave como el error cuadrático medio (RMSE), el error absoluto medio (MAE) y el coeficiente de determinación (R^2). Los resultados del estudio muestran que el modelo LSTM fue el más preciso, alcanzando un R^2 de 0,701, un RMSE de 0,087 y un MAE de 0,056. Esto lo convierte en la mejor opción para capturar dependencias temporales en los datos de series temporales complejas. En comparación, los modelos basados en transformadores demostraron tener potencial, pero no lograron la misma precisión que los modelos LSTM, especialmente en datos temporales más complicados. El estudio concluye que el LSTM es más eficaz en la predicción del ICA, equilibrando tanto la precisión como la eficiencia computacional, o que podría ayudar en intervenciones para mitigar la contaminación del aire.

Palabras clave: Índice de calidad del aire, RNN, LSTM, Transformadores, Pronóstico de contaminación.

^{1,*}Unidad académica de informática ciencias de la computación e innovación tecnológica, Universidad católica de Cuenca, Ecuador.

Corresponding author ✉: juan.pazmiño@ucacue.edu.ec.

Suggested citation: Buestán Andrade, P. A.; Carrión Zamora, P.E.; Chamba Lara, A.E. and Pazmiño Piedra, J.P. "A comprehensive evaluation of ai techniques for air quality index prediction: RNNs and transformers," *Ingenius, Revista de Ciencia y Tecnología*, N.º 33, pp. 60-75, 2025, DOI: <https://doi.org/10.17163/ings.n33.2025.06>.

1. Introduction

Air pollution poses a significant challenge to sustainable development due to its profound impact on public health, accounting for approximately 7 million deaths globally in 2019, according to the World Health Organization (WHO) [1, 2]. Despite the health benefits of clean air, a substantial portion of the population resides in urban areas or near industrial facilities with high levels of vehicular emissions [3]. The combustion of fossil fuels releases harmful pollutants, including carbon monoxide (CO), ozone (O_3), sulfur dioxide (SO_2), nitrogen dioxide (NO_2), and particulate matter ($PM_{2.5}$ and PM_{10}), which adversely affect human health and the environment [4]. The Air Quality Index (AQI) is a crucial metric for assessing and managing air quality, providing a comprehensive measure to evaluate pollution levels and their implications [5, 6].

The Air Quality Index (AQI) has been extensively studied for its environmental impacts [7–9], economic implications [7], [10] and predictive applications using data from monitoring stations [11–13]. Methods for AQI prediction are broadly classified into numerical and data-driven models [14]. Traditional statistical approaches, such as linear regression [15, 16] are employed alongside machine learning (ML) algorithms [17, 18] and hybrid models that integrate elements of both methodologies [14], [19]. Since the early 21st century, ML techniques, including artificial neural networks (ANN), support vector machines (SVM), extreme learning machines (ELM), and k-nearest neighbors (KNN), have become dominant in AQI prediction [20, 21]. Despite their widespread use, these methods exhibit limitations in processing temporal data, prompting the adoption of recurrent neural networks (RNNs), such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), for sequence prediction tasks.

Recent research has demonstrated the effectiveness of convolutional neural networks (CNNs) in AQI prediction. For instance, Yan et al. [22] developed models utilizing CNN, LSTM, and CNN-LSTM architectures, concluding that LSTM performs optimally for multi-hour forecasting, while CNN-LSTM is better suited for short-term predictions. Similarly, Hossain et al. [23] integrated GRU and LSTM for AQI forecasting in Bangladesh, achieving superior performance compared to individual techniques. To address long-term dependencies in sequential data, the Transformer model, which employs an encoder-decoder architecture, has emerged as a promising solution [24]. Guo et al. [25] applied a Transformer-based network, BERT, for AQI forecasting in Shanxi, China, achieving superior accuracy compared to LSTM. Ma et al. [26] developed the Informer model for AQI prediction in Yanan, China, demonstrating notable improvements in reliability and precision. Additionally, Xie et al. [27] proposed a parallel multi-input Transformer model for AQI forecasting

in Shanghai.

Comparing AI models for AQI prediction is critical due to variations in databases, evaluation metrics, and algorithms, which significantly influence model performance. Identifying the most effective model is essential for enhancing prediction accuracy and supporting informed decisions in air quality management, a key factor in public health and urban planning.

2. Materials and methods

2.1. General Description

The methodology of this research consists of the following steps, as illustrated in Figure 1:

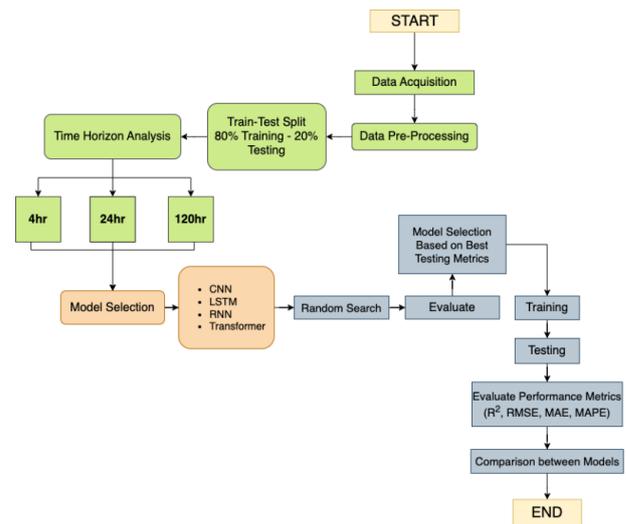


Figure 1. Research methodology flowchart

- Data acquisition and preprocessing.
- Splitting data into training (80%) and testing (20%) subsets.
- Analyzing the time horizon with options set at 4, 24, and 120 hours.
- Selecting models: CNN, LSTM, RNN, and Transformer architectures.
- Applying a random search for optimal hyperparameter selection.
- Evaluating models using the best metrics from tests.
- Training the selected models with the best hyperparameters.
- Testing the trained models.
- Evaluating model performance using R^2 , RMSE, MAE, and MAPE.

- Comparing results across models.
- Drawing conclusions based on the model comparison and overall analysis.

2.2. Study case

The time series data used in this study were obtained from a weather station in Cuenca, Azuay, Ecuador, located at 7-77 Bolivar Street and Borrero Street (coordinates: latitude -2.897, longitude -79.00). Managed by the Empresa Municipal de Movilidad, Tránsito de Transporte de la Municipalidad de Cuenca (EMOV-EP), this station provides publicly available data for personal, research, and governmental use. Positioned in a central area characterized by its commercial, tourist, colonial, and residential significance, this station is part of a network of three monitoring stations in Cuenca, as depicted in Figure 2.



Figure 2. EMOV-EP weather stations

2.3. Data preprocessing

The meteorological station recorded gas emissions, including CO , NO_2 , O_3 , $PM_{2.5}$, and SO_2 , at 10-minute intervals throughout 2022, generating approximately 52,560 records. These were exported in CSV format from the official EMOV-EP website. Where necessary, measurement units for pollutants were converted (e.g., CO from mg/m^3 to ppm) to facilitate AQI calculation. Hourly averages were computed, reducing the dataset to 8,760 records. After filtering out null values and irrelevant data, 7,425 records were retained for analysis. Table 1 provides an overview of the time series data, while Figure 3 illustrates the hourly variation of the recorded gases.

Table 1. Summary of the research time series

Statistics	CO	NO_2	O_3	$PM_{2.5}$	SO_2
Count	7425	7425	7425	7425	7425
Mean	9,909	0.258	0.275	15,040	0.038
Std	7,658	0.112	0.239	10,611	0.067
Min	0.024	0.040	0.000	0.100	0.000
25%	3,704	0.175	0.061	8,010	0.003
50%	8,946	0.241	0.223	12,100	0.011
75%	14,189	0.322	0.435	19,660	0.041
Max	50,454	0.811	1,147	75,330	0.629

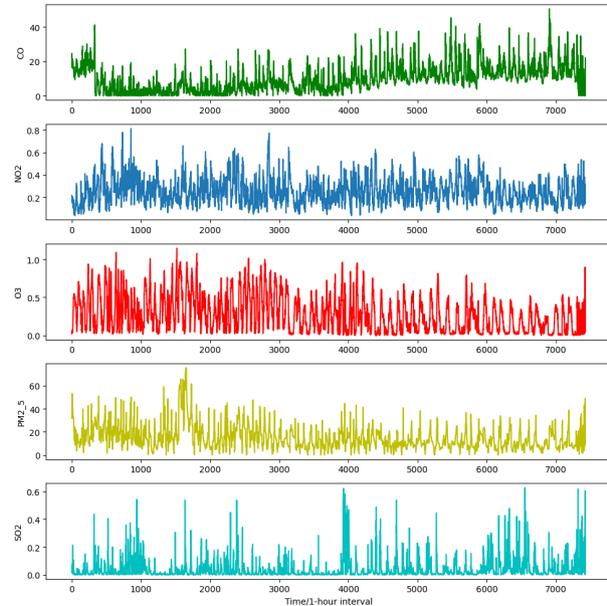


Figure 3. CO , NO_2 , O_3 , $PM_{2.5}$, and SO_2 registered per hour

Using the filtered values, the AQI for each pollutant was calculated following the guidelines outlined in the air quality report published by the U.S. Environmental Protection Agency [28]. The AQI computation is based on the pollutant concentration and is determined using Equation (1).

$$I_p = \frac{I_{HI} - I_{LO}}{BP_{HI} - BP_{LO}}(C_p - BP_{LO}) + I_{LO} \quad (1)$$

Where:

- I_p is the contaminant index p.
- C_p is the rounded concentration of the pollutant p.
- BP_{HI} is the cut-off point that is greater than or equal to C_p .
- BP_{LO} is the cutoff point that is less than or equal to C_p .
- I_{HI} is the AQI value corresponding to BP_{HI} .
- I_{LO} is the AQI value corresponding to BP_{LO} .

The individual contaminant index values (I_p) for each pollutant p are calculated independently, and the final AQI is determined by selecting the maximum value from the set of calculated indices. This selection methodology ensures that the final AQI reflects the pollutant exhibiting the highest potential for adverse health impacts, providing a comprehensive assessment of air quality conditions.

To analyze the relationship between the pollutants and the calculated AQI, a correlation matrix was constructed, as illustrated in Figure 4.

Figure 4 illustrates that, in this specific case, based solely on the pollutant data recorded by the meteorological station, the AQI value exhibits a stronger correlation with CO and NO_2 , while its correlation with O_3 , $PM_{2.5}$ and SO_2 is minimal. This information will be considered for the development and configuration of the AI models in subsequent analyses.

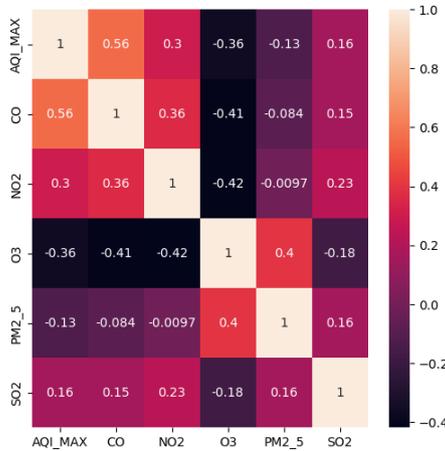


Figure 4. Correlation matrix of the different pollutants

Standardization involves transforming the data such that it has a mean of 0 and a standard deviation of 1. This process is analytically represented by Equation (2), which was utilized in this research:

$$X_{Standardization} = \frac{x - mean(x)}{standard\ deviation} \quad (2)$$

Where:

- X: It is the required value to be normalized.
- Mean: The arithmetic means of the distribution.
- Standard deviation: Standard deviation of the distribution.

Finally, after standardizing the time series values, the dataset was partitioned, with 80% allocated for training and 20% reserved for testing.

2.4. Sliding window

Hota, et al. [29] highlight that a commonly employed technique in time series analysis is the creation of sliding windows, which provides a temporal approximation of the true value of the time series data. This method accumulates historical time series data within a specified window to predict the subsequent value. Figure 5 illustrates the sliding window process with a window size of 5.

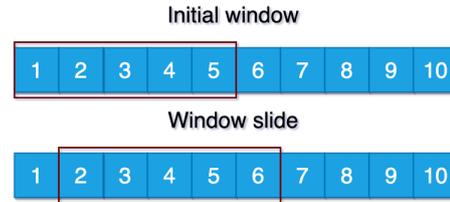


Figure 5. Sliding window process

Considering the aforementioned approach, sliding windows of 4, 24, and 120 hours were utilized to predict subsequent time intervals.

2.5. Deep learning

2.5.1. Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Long short-term memory (LSTM), Transformer model

Recurrent Neural Networks (RNNs) are a class of neural networks specifically designed for processing sequential data. Their architecture enables the output of one layer to loop back into the input, allowing the network to retain memory of prior states. This capability makes RNNs particularly effective for tasks requiring contextual or historical information, including time series prediction, natural language processing, and speech recognition [30]. A standard RNN configuration is illustrated in Figure 7.

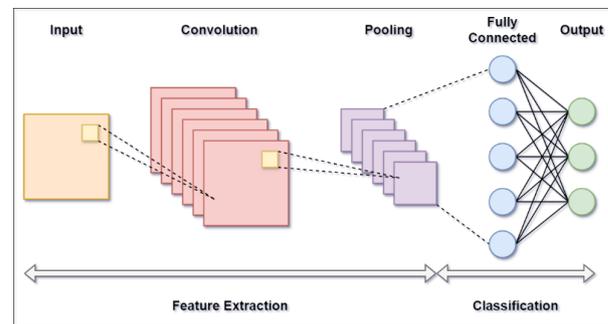


Figure 6. CNN standard configuration

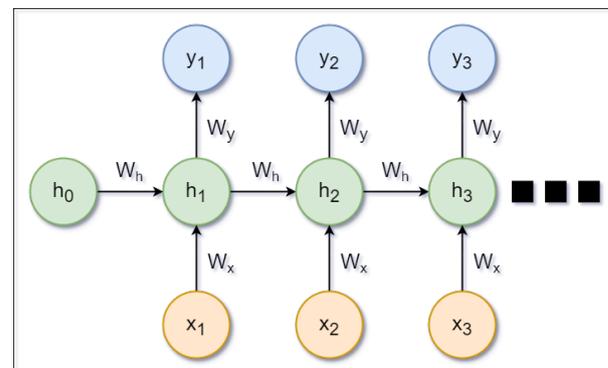


Figure 7. RNN standard configuration

Long Short-Term Memory Networks (LSTMs) were developed to address the limitations of traditional RNNs, such as the vanishing gradient problem, by incorporating a memory cell capable of retaining information over extended periods [31]. Each LSTM cell comprises three gates: the input gate, which controls the incorporation of new information; the forget gate, which eliminates irrelevant data; and the output gate, which determines the information to be passed to the next step [31]. A standard LSTM configuration is depicted in Figure 8.

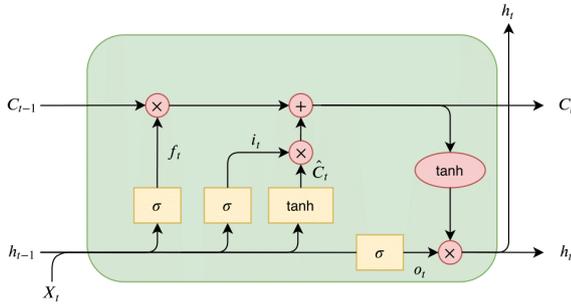


Figure 8. LSTM standard configuration [32]

Transformers have garnered considerable attention for their outstanding performance across various domains, including natural language processing (NLP), computer vision, and speech processing. Renowned for their ability to model long-term dependencies and complex interactions in sequential data, Transformers are particularly well-suited for time series prediction tasks [24]. The architecture implemented in this study is depicted in Figure 9.

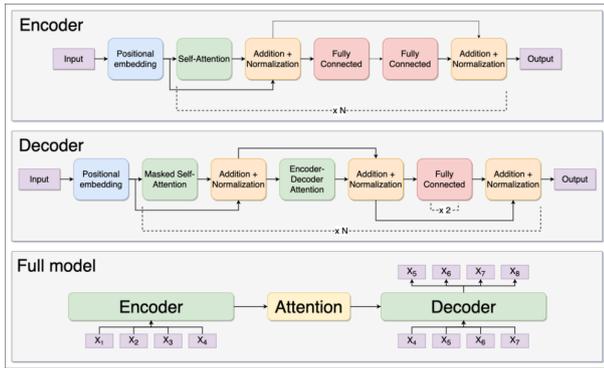


Figure 9. Transformers model architecture. Adapted from Youness [33]

2.6. Performance metrics

In the study conducted by Méndez, et al. [20], research aimed to identify the primary factors influencing air quality prediction during the period 2011–2021. The authors found that the most commonly applied metrics for evaluating machine learning (ML) models include

RMSE, MAE, MAPE, ACC, and R^2 as illustrated in Figure 10.

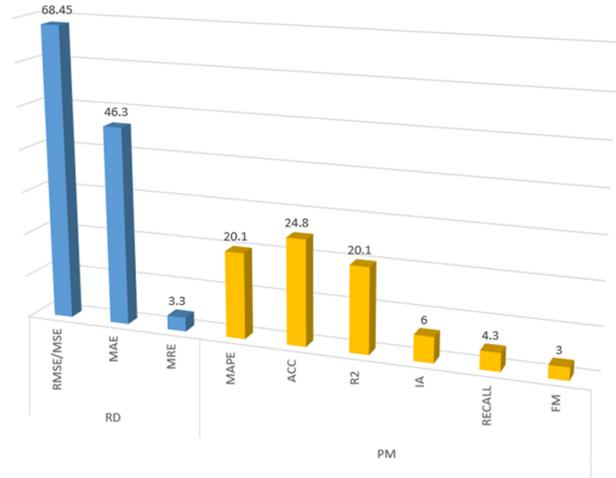


Figure 10. Evaluation of metrics usage [20]

2.6.1. Root mean square error (RMSE)

Root mean square error (RMSE) is a widely used metric that quantifies the average difference between predicted and observed values [34]. The formula for RMSE is presented in Equation (3):

$$RMSE = \sqrt{\frac{1}{N} \sum_{I=1}^N (y_I - Y_I)^2} \quad (3)$$

Where:

- n : number of samples.
- y_I : observed value.
- Y_I : predicted value.
- $(y_I - Y_I)^2$: squared error between predicted and observed values.

2.6.2. Mean absolute error (MAE)

Mean absolute error (MAE) is a metric used to assess the accuracy of a model by calculating the average of the absolute errors between predicted and observed values [34]. The formula for MAE is presented in Equation (4):

$$MAE = \frac{1}{n} \sum_{I=1}^N |y_I - Y_I| \quad (4)$$

Where:

- n : number of samples.
- y_I : observed value.

- y_I : predicted value.
- $|y_I - Y_I|$: absolute error between predicted and observed values

2.6.3. Mean absolute percentage error (MAPE)

Mean Absolute Percentage Error (MAPE) quantifies the average error as a percentage of the observed values, providing a scale-independent metric that facilitates comparisons across different models [34]. The formula for MAPE is presented in Equation (5):

$$MAPE = \frac{100\%}{N} \sum_{I=1}^N \left| \frac{y_I - Y_I}{y_i} \right| \quad (5)$$

Where:

- n: number of samples.
- y_I : observed value.
- y_I : predicted value.
- $\left| \frac{y_I - Y_I}{y_i} \right|$: relative absolute error, indicating the proportional deviation of the prediction from the actual value.

2.6.4. Coefficient of determination (R^2)

The coefficient of determination R^2 quantifies the proportion of variance in the dependent variable that is explained by the independent variables of the model. An R^2 value of 1 indicates that the model perfectly explains the variability in the data, whereas a value of 0 signifies that the model does not explain any variability [34]. The formula for R^2 is presented in Equation (6):

$$R^2 = \frac{\sum_{I=1}^N (y_I - Y_I)^2}{\sum_{I=1}^N (y_I - \bar{y})^2} \quad (6)$$

Where:

- n: number of samples.
- y_I : observed value.
- y_I : predicted value.
- \bar{y} : mean of all observed values y_I .
- $\sum_{I=1}^N (y_I - Y_I)^2$: sum of squared prediction errors.
- $\sum_{I=1}^N (y_I - \bar{y})^2$: sum of squared deviations of the observed values from their mean.

2.7. Random search

In ML, random search (RS) is an optimization technique used to identify optimal hyperparameters by exploring random combinations within a predefined parameter space. This approach is more efficient and computationally less expensive compared to exhaustive search methods [35]. Figure 11 illustrates the sequence of steps involved in the RS process.

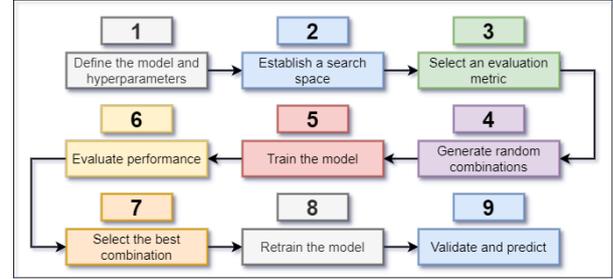


Figure 11. Methodology of the random search technique

The process involves the following steps:

- Identifying the AI model to optimize and define its hyperparameters.
- Setting ranges for each hyperparameter.
- Specifying evaluation metrics, such as R^2 , MAE, and RMSE.
- Randomly selecting combinations of hyperparameters.
- Training the model and evaluating its performance.
- Choosing the best-performing combination and retraining the model.
- Validating the performance of the optimized model.

3. Results and discussion

3.1. Correlation Analysis

Figure 4 illustrates the relationship between variables using Pearson’s Correlation Coefficient to assess their correlation with maximum AQI concentration (AQI_{MAX}). The heatmap reveals that CO has the strongest positive correlation with AQI_{MAX} (0.56), followed by NO_2 (0.3) and SO_2 (0.16), while O_3 shows a negative correlation (-0.36). Furthermore, CO is moderately correlated with NO_2 (0.36) and SO_2 (0.15), while O_3 demonstrates inverse correlations with both CO (-0.41) and NO_2 (-0.42). $PM_{2.5}$ exhibits weak negative correlations with AQI_{MAX} , CO , and NO_2 . These results highlight CO as the variable most strongly correlated with AQI_{MAX} , providing critical guidance for selecting the input combinations outlined in Table 2.

Table 2. Combinations of input variables

Combination	Input variables
C1	CO, NO_2
C2	CO, O_3
C3	CO, $PM_{2.5}$
C4	CO, SO_2
C5	NO_2 , O_3
C6	NO_2 , $PM_{2.5}$
C7	NO_2 , SO_2
C8	O_3 , $PM_{2.5}$
C9	O_3 , SO_2
C10	CO, NO_2 , O_3
C11	CO, NO_2 , O_3
C12	CO, NO_2 , $PM_{2.5}$
C13	CO, NO_2 , SO_2
C14	CO, O_3 , $PM_{2.5}$
C15	CO, O_3 , SO_2
C16	CO, $PM_{2.5}$, SO_2
C17	NO_2 , O_3 , $PM_{2.5}$
C18	NO_2 , O_3 , SO_2
C19	NO_2 , $PM_{2.5}$, SO_2
C20	O_3 , $PM_{2.5}$, SO_2
C21	CO, NO_2 , O_3 , $PM_{2.5}$
C22	CO, NO_2 , O_3 , SO_2
C23	CO, NO_2 , $PM_{2.5}$, SO_2
C24	CO, O_3 , $PM_{2.5}$, SO_2
C25	NO_2 , O_3 , $PM_{2.5}$, SO_2
C26	CO, NO_2 , O_3 , $PM_{2.5}$, SO_2

3.2. Results of the AI models

This section analyzes the performance outcomes of each model, considering the input combinations outlined in Table 3, the type of AI model employed, and the associated evaluation metrics.

Table 3. CNN model results

IP	BS	RMSE	MAE	R^2	MAPE (%)	PPS
C1	24	0.08	0.06	0.68	16.71	35805.01
C2	24	0.09	0.06	0.68	16.82	58737.06
C3	24	0.09	0.06	0.67	17.26	49748.10
C4	24	0.09	0.06	0.68	17.20	39398.15
C5	24	0.09	0.07	0.65	18.52	43885.02
C6	24	0.10	0.07	0.64	17.65	65236.56
C7	24	0.10	0.07	0.62	16.76	65931.50
C8	24	0.09	0.06	0.69	17.29	53651.60
C9	24	0.09	0.07	0.66	17.88	59559.18
C10	24	0.10	0.07	0.64	17.81	36020.17
C11	24	0.09	0.07	0.65	16.84	44975.06
C12	24	0.09	0.07	0.67	17.59	73533.63
C13	24	0.09	0.06	0.67	16.24	58225.91
C14	24	0.08	0.06	0.69	15.48	49661.10
C15	24	0.09	0.06	0.68	16.42	56290.50
C16	24	0.09	0.07	0.66	17.82	35025.14
C17	24	0.09	0.06	0.68	16.85	76968.79
C18	24	0.09	0.07	0.67	17.31	73779.83
C19	24	0.10	0.07	0.63	18.30	65446.41
C20	24	0.09	0.06	0.67	18.82	76016.82
C21	24	0.09	0.07	0.66	16.72	59341.83
C22	24	0.09	0.06	0.68	16.31	67292.28
C23	24	0.09	0.06	0.68	16.85	44773.68
C24	24	0.09	0.06	0.67	16.73	79901.70
C25	24	0.09	0.06	0.68	16.71	55093.58
C26	24	0.09	0.07	0.67	16.58	75801.26

3.2.1. CNN results

The CNN model was evaluated using 26 different input parameter combinations, all based on a 24-hour sliding window (Table 4). The R^2 ranged from 0.620 to 0.698, with the combination of CO, O_3 , $PM_{2.5}$, and SO_2 (Row 24) achieving the highest performance at an R^2 value of 0.698. Conversely, the combination involving NO_2 and SO_2 (Row 7) yielded the lowest performance, with an R^2 of 0.620. Regarding error metrics, the RMSE ranged from 0.087 to 0.097, while the MAE varied from 0.060 to 0.068. Notably, the combination of CO, O_3 , $PM_{2.5}$, and SO_2 also exhibited the lowest errors, further highlighting its superior performance.

Table 4. RNN model results

IP	BS	RMSE	MAE	R^2	MAPE(%)	PPS
C1	24	0.09	0.06	0.56	21.97	20477.85
C2	24	0.09	0.06	0.57	21.32	39843.79
C3	24	0.09	0.06	0.57	20.65	41488.60
C4	24	0.09	0.06	0.56	20.48	40265.98
C5	24	0.09	0.06	0.56	22.92	35324.96
C6	24	0.09	0.06	0.56	21.41	37750.82
C7	24	0.09	0.06	0.55	21.36	35319.93
C8	24	0.10	0.06	0.55	21.21	31675.68
C9	24	0.09	0.06	0.56	21.51	34623.43
C10	24	0.09	0.06	0.57	20.98	41072.63
C11	24	0.09	0.06	0.56	21.73	37360.15
C12	24	0.09	0.06	0.57	21.00	41908.50
C13	24	0.09	0.06	0.56	21.00	34451.12
C14	24	0.09	0.06	0.56	21.73	40667.03
C15	24	0.09	0.06	0.56	21.31	36455.38
C16	24	0.09	0.06	0.56	20.74	34073.62
C17	24	0.09	0.06	0.55	21.55	36014.80
C18	24	0.09	0.06	0.56	21.48	34888.69
C19	24	0.09	0.06	0.56	20.92	39017.72
C20	24	0.09	0.06	0.56	21.53	33701.56
C21	24	0.09	0.06	0.58	21.29	35349.87
C22	24	0.09	0.06	0.57	21.14	40265.69
C23	24	0.09	0.06	0.57	20.82	34805.23
C24	24	0.09	0.06	0.56	21.07	37740.95
C25	24	0.09	0.06	0.57	21.67	35611.69
C26	24	0.09	0.06	0.57	21.85	43654.79

The MAPE values, indicating prediction accuracy, ranged from 15.48% to 18.83%. Combinations involving CO and O_3 exhibited the lowest MAPE values, reflecting higher prediction accuracy. Conversely, combinations including NO_2 and $PM_{2.5}$ had higher MAPE values, indicating lower prediction accuracy. Computational efficiency varied significantly, with prediction rates ranging from 35,805 to 79,091 predictions per second. More complex combinations, such as CO, NO_2 , O_3 , $PM_{2.5}$, and SO_2 (Row 26), required up to 4.5 GB of RAM, yet achieved superior prediction rates.

In summary, the combination of CO, O_3 , $PM_{2.5}$, and SO_2 (Row 24) emerged as the most accurate, achieving the highest R^2 , the lowest errors, and strong computational efficiency. In contrast, combinations including NO_2 and SO_2 underperformed across all metrics, suggesting these variables have a lesser impact on prediction accuracy.

Where

- IP: Input Parameters
- BS: Best Window
- PPS: Predictions Per Second

3.2.2. RNN results

The RNN model was evaluated using 26 different combinations of input parameters within a 24-hour sliding window (Table 5). The R^2 values ranged from 0.533 to 0.576, with the combination of CO , NO_2 , O_3 , and $PM_{2.5}$ (Row 21) achieving the highest R^2 of 0.576. Conversely, the combination involving NO_2 and SO_2 (Row 7) showed the lowest performance, with an R^2 of 0.533. The RMSE values ranged from 0.092 to 0.097, while the MAE ranged from 0.058 to 0.063, with the combination of CO , NO_2 , O_3 , and $PM_{2.5}$, demonstrating the lowest error rates, indicating superior performance.

Table 5. LSTM model results

IP	BS	RMSE	MAE	R^2	MAPE (%)	PPS
C1	24	0.09	2253.00	0.69	15.65	7049.54
C2	24	0.09	0.06	0.69	15.90	8122.33
C3	720	0.09	0.06	0.69	16.26	147.46
C4	720	0.09	0.06	0.67	16.39	141.67
C5	144	0.09	0.06	0.69	16.34	3336.96
C6	144	0.09	0.06	0.68	16.37	3031.92
C7	144	0.09	0.06	0.69	15.31	2987.68
C8	24	0.09	0.06	0.68	16.23	6412.71
C9	24	0.09	0.06	0.68	15.57	6720.89
C10	24	0.09	0.06	0.67	16.11	1544.31
C11	144	0.09	0.06	0.69	16.83	6513.84
C12	144	0.09	0.06	0.69	16.08	869.33
C13	144	0.09	0.06	0.70	15.35	806.71
C14	24	0.09	0.06	0.68	16.09	8009.40
C15	144	0.09	0.06	0.69	15.73	823.46
C16	144	0.09	0.06	0.70	16.50	3871.16
C17	720	0.09	0.06	0.68	16.87	706.77
C18	24	0.09	0.06	0.67	16.36	1023.41
C19	24	0.09	0.06	0.68	15.84	8613.76
C20	24	0.09	0.06	0.68	16.18	8341.78
C21	24	0.09	0.06	0.69	16.50	1036.17
C22	24	0.09	0.06	0.68	16.17	1086.39
C23	144	0.09	0.06	0.70	16.08	501.98
C24	720	0.09	0.06	0.69	16.72	116.44
C25	24	0.09	0.06	0.69	17.31	944.15
C26	24	0.09	0.06	0.70	16.00	6525.65

The MAPE values ranged between 20.48% and 22.92%, with the lowest values observed in combinations that included CO and $PM_{2.5}$, suggesting better prediction accuracy. In contrast, combinations involving NO_2 and O_3 exhibited higher MAPE values, indicating lower accuracy. Computational efficiency varied significantly, with prediction rates spanning 20,478 to 43,654 predictions per second. More complex combinations, such as CO , NO_2 , O_3 , $PM_{2.5}$, and SO_2 (Row 26), required up to 23.4 GB of RAM but demonstrated higher prediction throughput.

In conclusion, the combination of CO , NO_2 , O_3 , and $PM_{2.5}$ emerged as the most accurate, achieving the highest R^2 , the lowest error metrics, and strong computational efficiency. In contrast, combinations including NO_2 and SO_2 consistently underperformed across all metrics, suggesting these variables have a limited impact on prediction accuracy.

3.2.3. LSTM results

The LSTM model was evaluated using 26 different combinations of input parameters, primarily with a 24-hour sliding window, although some configurations employed 144-hour or 720-hour windows (Table 6). The R^2 values ranged from 0.669 to 0.701, with the combination of CO , NO_2 , $PM_{2.5}$, and SO_2 (Row 23) achieving the highest R^2 of 0.701. Conversely, the combination of NO_2 , O_3 , and SO_2 (Row 18) exhibited the lowest R^2 at 0.669. The RMSE values varied between 0.087 and 0.092, while the MAE ranged from 0.056 to 0.062, with the combination of CO , NO_2 , $PM_{2.5}$, and SO_2 demonstrating the lowest error metrics, signifying superior performance.

Table 6. Transformers model results

IP	BS	RMSE	MAE	R^2	MAPE (%)	PPS
C1	24	0.09	2253.00	0.69	15.65	7049.54
C2	24	0.09	0.06	0.69	15.90	8122.33
C3	720	0.09	0.06	0.69	16.26	147.46
C4	720	0.09	0.06	0.67	16.39	141.67
C5	144	0.09	0.06	0.69	16.34	3336.96
C6	144	0.09	0.06	0.68	16.37	3031.92
C7	144	0.09	0.06	0.69	15.31	2987.68
C8	24	0.09	0.06	0.68	16.23	6412.71
C9	24	0.09	0.06	0.68	15.57	6720.89
C10	24	0.09	0.06	0.67	16.11	1544.31
C11	144	0.09	0.06	0.69	16.83	6513.84
C12	144	0.09	0.06	0.69	16.08	869.33
C13	144	0.09	0.06	0.70	15.35	806.71
C14	24	0.09	0.06	0.68	16.09	8009.40
C15	144	0.09	0.06	0.69	15.73	823.46
C16	144	0.09	0.06	0.70	16.50	3871.16
C17	720	0.09	0.06	0.68	16.87	706.77
C18	24	0.09	0.06	0.67	16.36	1023.41
C19	24	0.09	0.06	0.68	15.84	8613.76
C20	24	0.09	0.06	0.68	16.18	8341.78
C21	24	0.09	0.06	0.69	16.50	1036.17
C22	24	0.09	0.06	0.68	16.17	1086.39
C23	144	0.09	0.06	0.70	16.08	501.98
C24	720	0.09	0.06	0.69	16.72	116.44
C25	24	0.09	0.06	0.69	17.31	944.15
C26	24	0.09	0.06	0.70	16.00	6525.65

The MAPE values ranged from 15.31% to 17.31%, with the lowest values observed in combinations including CO and NO_2 , indicating superior prediction accuracy. Conversely, combinations involving NO_2 and $PM_{2.5}$ exhibited higher MAPE values, suggesting lower accuracy. Computational efficiency varied significantly, with predictions per second ranging from 116 to 8,613. Shorter sliding windows generally resulted in higher prediction rates but required increased RAM usage. For instance, complex combinations such as CO , NO_2 ,

O_3 , $PM_{2.5}$, and SO_2 (Row 26) demanded up to 4.4 GB of RAM and achieved moderate prediction rates. In contrast, simpler combinations like CO and NO_2 required less RAM (3.3 GB) but exhibited lower prediction rates.

In summary, the combination of CO , NO_2 , $PM_{2.5}$, and SO_2 (Row 23) emerged as the most accurate, achieving the highest R^2 and the error metrics, albeit with increased computational demands. Conversely, combinations involving NO_2 and SO_2 exhibited consistently lower performance across all evaluated metrics.

3.2.4. Transformer results

In terms of error metrics, the RMSE ranged from 0.094 to 0.130, and the MAE from 0.068 to 0.099, with the lowest errors observed for the combination of CO , NO_2 , $PM_{2.5}$, and SO_2 . Higher errors, particularly in MAE, were noted for combinations involving CO and SO_2 , suggesting these inputs are less effective for accurate predictions. The MAPE values varied between 18.19% and 26.83%, with the lowest values associated with combinations involving O_3 and $PM_{2.5}$, while higher MAPE values were observed for combinations including CO and O_3 (Row 2).

Regarding computational efficiency, predictions per second ranged from 2,974 to 21,030. More complex combinations with a greater number of input parameters and longer sliding windows required higher RAM usage (up to 3.7 GB) but achieved faster prediction rates. Conversely, simpler combinations, such as CO and NO_2 required less RAM (1.8 GB) but exhibited slower prediction rates. Overall, the combination of CO , NO_2 , $PM_{2.5}$, and SO_2 (Row 23) emerged as the most accurate, achieving the highest R^2 and the lowest error metrics, albeit with increased computational requirements. In contrast, combinations involving NO_2 and O_3 showed inferior performance across all metrics.

3.2.5. AI model results analysis

When comparing the performance of the RNN, CNN, LSTM, and Transformer models, the Transformer model exhibits notably lower accuracy, with R^2 values ranging from 0.322 to 0.640. This broad range underscores significant challenges in capturing the variability of the output data, particularly when utilizing variables such as NO_2 and O_3 . While the Transformer model demonstrates higher computational efficiency—requiring 1.8 GB to 3.7 GB of RAM and achieving prediction rates between 2,974 and 21,030 predictions per second—this efficiency does not offset its lower predictive accuracy. As illustrated in Figure 12, the scatter plots for the Transformer model reveal substantial dispersion around the reference line, particularly at extreme AQI values. This deviation highlights the model's unreliability in these situations.

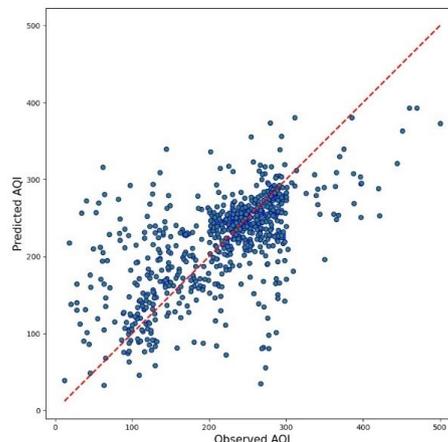


Figure 12. Transformer model predictions combination 23 (Best R^2 and RMSE)

The RNN model, while more accurate than the Transformer, demonstrates intermediate performance with R^2 values ranging from 0.533 to 0.576. The RNN achieves acceptable accuracy, with RMSE values between 0.092 and 0.097 and MAE values ranging from 0.058 to 0.063. However, its MAPE values, which fall between 20.48% and 22.92%, are higher than those observed for the CNN and LSTM models. As depicted in Figure 13, the RNN scatter plots exhibit a higher density of points near the $y = x$ line compared to the Transformer model, suggesting improved overall alignment with observed values. Nonetheless, significant dispersion persists at the extreme AQI levels, highlighting variability in accuracy when predicting high or low AQI values.

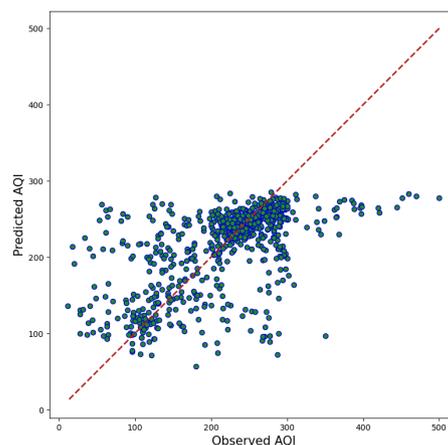


Figure 13. Transformer model predictions combination 10 (Average R^2)

In terms of computational efficiency, the RNN exhibits a balanced performance, with RAM usage ranging from 22.8 GB to 23.4 GB and prediction rates between 20,478 and 43,654 predictions per second. While not exceptional, this performance positions the RNN as a viable option, offering a reasonable trade-off between accuracy and efficiency.

The CNN model exhibited a higher R^2 range of 0.620 to 0.698, demonstrating its strong capability to capture data variability and provide accurate predictions. Among the tested combinations, CO , O_3 , $PM_{2.5}$, and SO_2 yielded the best performance. The errors were relatively low, with RMSE values between 0.087 and 0.097 and MAE values ranging from 0.060 to 0.068. Additionally, MAPE values between 15.31% and 18.83% further underscored the model’s predictive accuracy. As illustrated in Figure 14, the scatter plots reveal the CNN model’s ability to deliver consistent predictions, with a significant clustering of data points around the regression line, even at extreme values, thereby minimizing errors. Although the CNN model is less computationally efficient than the Transformer, it maintains a reasonable balance, with RAM usage ranging from 3.8 GB to 4.5 GB and a prediction rate of 35,805 to 79,091 predictions per second.

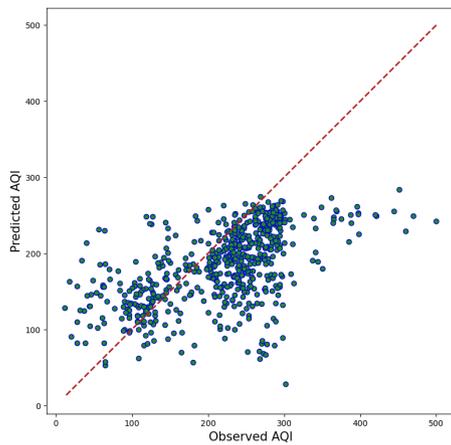


Figure 14. Transformer model predictions combination 5 (Worst R^2)

The LSTM model demonstrates superior accuracy among the evaluated models, with an R^2 range of 0.669 to 0.701, highlighting its exceptional ability to discern patterns within the data. It outperforms other models in error metrics, achieving RMSE values between 0.087 and 0.092 and MAE values ranging from 0.056 to 0.062. Additionally, the LSTM model exhibits the lowest MAPE values, ranging from 15.31% to 17.31%, underscoring its remarkable prediction accuracy. As illustrated in Figure 15, the scatter plots for the LSTM model reveal a high concentration of data points near the reference line, with minimal scatter, even for extreme AQI values. This consistency and precision position the LSTM as a robust and reliable option for applications where accuracy is critical.

However, the LSTM model is the least computationally efficient among those evaluated, with RAM usage ranging from 3.3 GB to 4.5 GB and prediction speeds between 116 and 8,613 predictions per second. This relative inefficiency, especially when using longer

sliding windows, may limit its applicability in scenarios where processing speed is critical. Nonetheless, its exceptional accuracy establishes it as a highly reliable option for applications where precision takes precedence over computational efficiency.

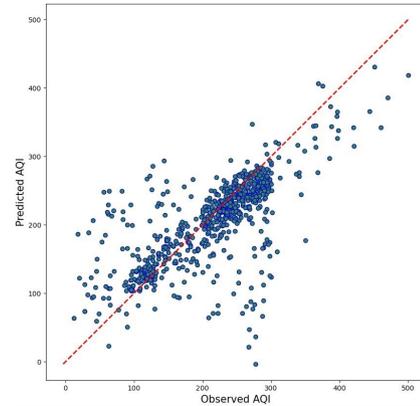


Figure 15. RNN model predictions combination 21 (Best R^2)

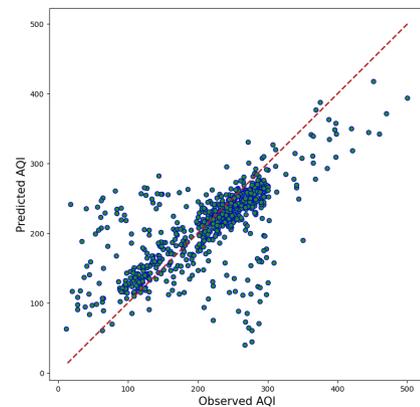


Figure 16. RNN model predictions combination 20 (Average R^2)

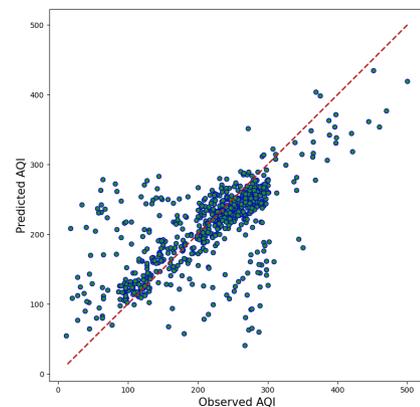


Figure 17. RNN model predictions combination 7 (Worst R^2)

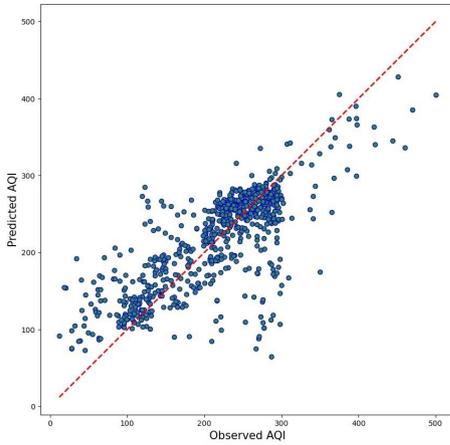


Figure 18. CNN model prediction combination 24 (Best R^2)

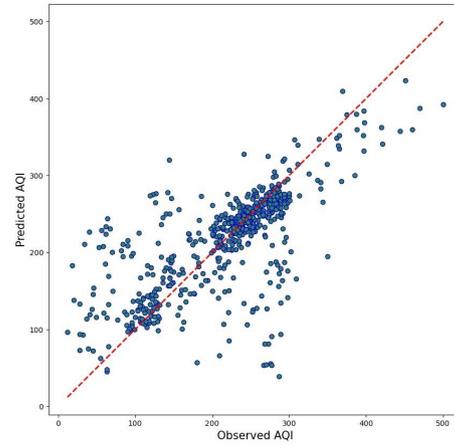


Figure 21. LSTM model prediction combination 23 (Best R^2)

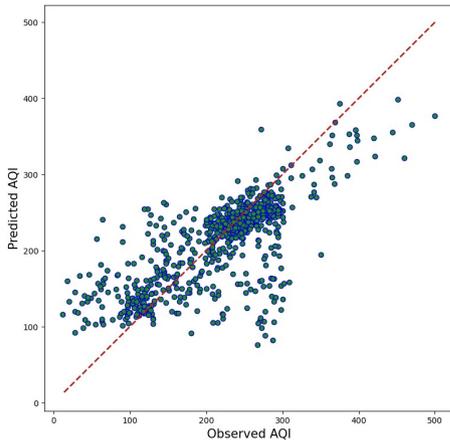


Figure 19. CNN model prediction combination 18 (Average R^2)

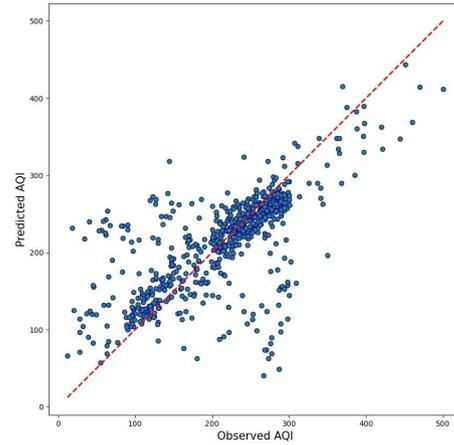


Figure 22. LSTM model prediction combination 9 (Best R^2)

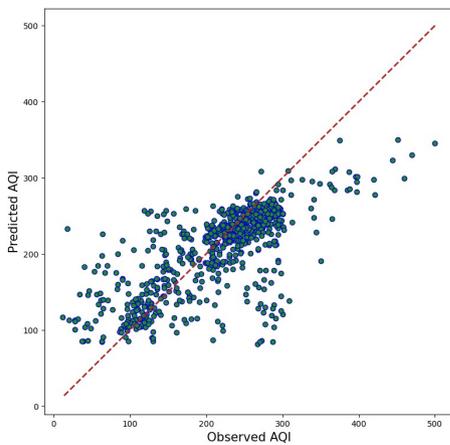


Figure 20. CNN model prediction combination 7 (Worst R^2)

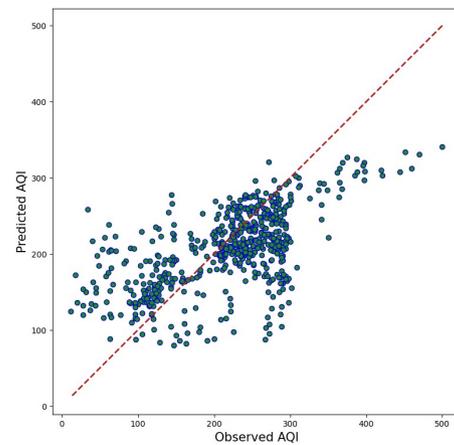


Figure 23. LSTM model prediction combination 14 (Average R^2)

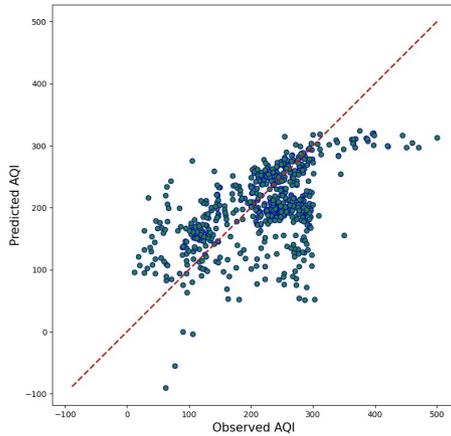


Figure 24. LSTM model prediction combination 18 (Worst R^2)

3.2.6. Comparison with related studies

This study evaluates the effectiveness of LSTM, GRU, RNN, CNN and Transformer models for predicting the Air Quality Index (AQI) in Cuenca, Ecuador, and compares the results with findings from other relevant research, such as the study conducted by Cui et al. [36], which focused on $PM_{2.5}$ prediction using Transformer and CNN-LSTM-Attention models in Beijing, China.

A key distinction between the studies lies in the datasets and their characteristics. This study utilizes data from a single meteorological station in Cuenca, comprising 7,425 records for the year 2022. In contrast, the Beijing study employs data from 12 monitoring stations collected over four years (2013–2017), totaling over 35,000 records. The richer dataset in Beijing allowed researchers to incorporate seasonal variations and long-term dependencies, critical factors for accurate $PM_{2.5}$ prediction.

Additionally, their Transformer model was enhanced with multi-head attention mechanisms and positional encoding, enabling more effective capture of complex temporal patterns and seasonal fluctuations [36].

Regarding model performance, this study demonstrates that the LSTM model achieves the highest accuracy for AQI prediction in Cuenca, with an R^2 of 0.701, surpassing the Transformer model's performance, which achieved an approximate R^2 of 0.68. Conversely, in the Beijing study, the Transformer model significantly outperformed the CNN-LSTM-Attention architecture, achieving an R^2 of 0.944 compared to 0.836. This superior performance was attributed to the Transformer's capacity to handle both abrupt meteorological changes and long-term trends, particularly during complex seasonal transitions such as those observed in autumn and winter [36].

Another notable difference is the prediction horizon. This study evaluated short- to medium-term prediction

windows (4, 24, and 120 hours), whereas the Beijing study focused on hourly predictions. The Transformer model in Beijing proved particularly effective in capturing sudden pollutant variations driven by meteorological changes, underscoring its suitability for high-frequency predictions in dynamic environments [36].

These findings underscore the necessity of tailoring AI architectures to the unique characteristics of specific datasets and prediction objectives. Future research should focus on developing hybrid models that leverage the complementary strengths of LSTM and Transformer architectures, aiming to effectively tackle both local and regional air quality forecasting challenges.

4. Conclusions

In this study, the performance of various artificial intelligence models, including RNN, CNN, LSTM, and Transformers, was evaluated and compared for the task of predicting the Air Quality Index (AQI). The findings reveal that the LSTM model consistently outperformed the other models, achieving an R^2 of 0.701 and an RMSE of 0.087. Its superior performance was particularly evident when using variable combinations such as CO , NO_2 , $PM_{2.5}$, and SO_2 . The analysis underscores the LSTM model's effectiveness in capturing complex temporal relationships among these variables, establishing it as a reliable and valuable tool for accurate AQI prediction in diverse scenarios.

Although Transformers models have demonstrated exceptional performance in various fields such as natural language processing (NLP) and computer vision (CV), their application to AQI prediction, particularly with this specific dataset, reveals significant limitations. In this study, Transformers exhibited notable variability in performance, with coefficients of determination ranging from 0.322 to 0.640. These findings suggest that Transformers face challenges in effectively capturing the intrinsic complexity of the analyzed time-series data, particularly when using variable combinations that include NO_2 and O_3 . Despite their computational efficiency, the predictive accuracy of Transformers for AQI falls short compared to more competitive models such as LSTM.

In terms of computational efficiency, both CNN and LSTM models have demonstrated their suitability for real-time applications, offering an effective balance between accuracy and resource utilization. The LSTM model, in particular, stands out for its exceptional predictive accuracy, efficient RAM usage of approximately 4.4 GB, and its capability to perform a substantial number of predictions per second. This combination of high performance and resource efficiency makes LSTM especially well-suited for air quality prediction systems that require fast and precise responses, such as

real-time environmental monitoring applications.

Despite achieving satisfactory R^2 values, the analysis revealed considerable data dispersion and elevated error metrics across all implementations, rendering some models unsuitable for reliable deployment. This study serves as a comparative evaluation of artificial intelligence approaches, highlighting both the strengths and limitations of current AI architectures for AQI forecasting. The findings underscore the need for refinement in existing implementations while emphasizing the substantial potential of AI models for improving air quality predictions in the future.

In conclusion, the training of all models encountered limitations due to the relatively small size of the Cuenca dataset and the low correlation observed between certain pollutants. For future research, we recommend leveraging larger datasets and extending the analysis over longer periods to enhance model performance and generate more robust insights.

4.1. Future directions

The selection of Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), Transformer, and Recurrent Neural Networks (RNN) architectures as foundational models in this study was based on their established prominence in time series analysis and their significant contributions to advancements in AI research. Building upon these foundational architectures, future research should explore the incorporation of enhanced Transformer-based models through the integration of multi-head attention mechanisms and positional encoding schemes. These enhancements could enable more sophisticated modeling of temporal dependencies in air quality patterns. The implementation of Physics-Informed Neural Networks (PINNs) also presents a promising direction, as these architectures explicitly integrate fundamental atmospheric physics and chemical transport equations into the neural network framework. This approach offers the potential to bridge the gap between data-driven methodologies and theoretical models, enhancing the interpretability and accuracy of predictions. While established architectures such as Transformers, CNNs, RNNs, and LSTMs have demonstrated notable efficacy, exploring emerging methodologies like Neural Ordinary Differential Equations (Neural ODEs), Temporal Fusion Transformers, and Informer networks could yield even greater predictive capabilities. These novel approaches, though less widely adopted in AI research, may address existing challenges in modeling non-linear atmospheric dynamics and complex inter-variable correlations, thus advancing AQI forecasting to new levels of precision and reliability.

References

- [1] J. Zhang and S. Li, “Air quality index forecast in beijing based on cnn-lstm multi-model,” *Chemosphere*, vol. 308, p. 136180, 2022. [Online]. Available: <https://doi.org/10.1016/j.chemosphere.2022.136180>
- [2] M. Ansari and M. Alam, “An intelligent iot-cloud-based air pollution forecasting model using univariate time-series analysis,” *Arabian Journal for Science and Engineering*, vol. 49, no. 3, pp. 3135–3162, Mar 2024. [Online]. Available: <https://doi.org/10.1007/s13369-023-07876-9>
- [3] F. Cassano, A. Casale, P. Regina, L. Spadafina, and P. Sekulic, “A recurrent neural network approach to improve the air quality index prediction,” in *Ambient Intelligence – Software and Applications – 10th International Symposium on Ambient Intelligence*, P. Novais, J. Lloret, P. Chamoso, D. Carneiro, E. Navarro, and S. Omatu, Eds. Cham: Springer International Publishing, 2020, pp. 36–44. [Online]. Available: https://doi.org/10.1007/978-3-030-24097-4_5
- [4] K. Oguz and M. A. Pekin, *Prediction of Air Pollution with Machine Learning Algorithms*. 1–12: Turkish Journal of Science and Technology, 2024, vol. 19, no. 1. [Online]. Available: <https://doi.org/10.55525/tjst.1224661>
- [5] H. Jalali, F. Keynia, F. Amirteimoury, and A. Heydari, “A short-term air pollutant concentration forecasting method based on a hybrid neural network and metaheuristic optimization algorithms,” *Sustainability*, vol. 16, no. 11, 2024. [Online]. Available: <https://doi.org/10.3390/su16114829>
- [6] R. Patil, D. Dinde, and S. Powar, “A literature review on prediction of air quality index and forecasting ambient air pollutants using machine learning algorithms,” *International Journal of Innovative Science and Research Technology*, vol. 5, pp. 1148–1152, 09 2020. [Online]. Available: <http://dx.doi.org/10.38124/IJISRT20AUG683>
- [7] R. Hassan, M. Rahman, and A. Hamdan, “Assessment of air quality index (aqi) in riyadh, saudi arabia,” *IOP Conference Series: Earth and Environmental Science*, vol. 1026, no. 1, p. 012003, may 2022. [Online]. Available: <https://dx.doi.org/10.1088/1755-1315/1026/1/012003>
- [8] W.-T. Tsai and Y.-Q. Lin, “Trend analysis of air quality index (aqi) and greenhouse gas (ghg) emissions in taiwan and their regulatory countermeasures,” *Environments*, vol. 8, no. 4, 2021. [Online]. Available: <https://doi.org/10.3390/environments8040029>

- [9] G. Ravindiran, S. Rajamanickam, K. Kanagarathinam, G. Hayder, G. Janardhan, P. Arunkumar, S. Arunachalam, A. A. AlObaid, I. Warad, and S. K. Muniasamy, "Impact of air pollutants on climate change and prediction of air quality index using machine learning models," *Environmental Research*, vol. 239, p. 117354, 2023. [Online]. Available: <https://doi.org/10.1016/j.envres.2023.117354>
- [10] N. P. Canh, W. Hao, and U. Wongchoti, "The impact of economic and financial activities on air quality: a chinese city perspective," *Environmental Science and Pollution Research*, vol. 28, no. 7, pp. 8662–8680, Feb 2021. [Online]. Available: <https://doi.org/10.1007/s11356-020-11227-8>
- [11] T. Madan, S. Sagar, and D. Virmani, "Air quality prediction using machine learning algorithms –a review," in *2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, 2020, pp. 140–145. [Online]. Available: <https://doi.org/10.1109/ICACCCN51052.2020.9362912>
- [12] M. Kaur, D. Singh, M. Y. Jabarulla, V. Kumar, J. Kang, and H.-N. Lee, "Computational deep air quality prediction techniques: a systematic review," *Artificial Intelligence Review*, vol. 56, no. 2, pp. 2053–2098, Nov 2023. [Online]. Available: <https://doi.org/10.1007/s10462-023-10570-9>
- [13] S. Senthivel and M. Chidambaranathan, "Machine learning approaches used for air quality forecast: A review," *Revue d'Intelligence Artificielle*, vol. 36, no. 1, pp. 73–78, 2022. [Online]. Available: <https://doi.org/10.18280/ria.360108>
- [14] Q. Wu and H. Lin, "A novel optimal-hybrid model for daily air quality index prediction considering air pollutant factors," *Science of The Total Environment*, vol. 683, pp. 808–821, 2019. [Online]. Available: <https://doi.org/10.1016/j.scitotenv.2019.05.288>
- [15] Y. Xu, H. Liu, and Z. Duan, "A novel hybrid model for multi-step daily aqi forecasting driven by air pollution big data," *Air Quality, Atmosphere & Health*, vol. 13, no. 2, pp. 197–207, Feb 2020. [Online]. Available: <https://doi.org/10.1007/s11869-020-00795-w>
- [16] Z. Zhao, J. Wu, F. Cai, S. Zhang, and Y.-G. Wang, "A statistical learning framework for spatial-temporal feature selection and application to air quality index forecasting," *Ecological Indicators*, vol. 144, p. 109416, 2022. [Online]. Available: <https://doi.org/10.1016/j.ecolind.2022.109416>
- [17] J. A. Moscoso-López, D. Urda, J. González-Enrique, J. J. Ruiz-Aguilar, and I. Turias, "Hourly air quality index (aqi) forecasting using machine learning methods," in *15th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2020)*, 2021, pp. 123–132. [Online]. Available: <https://upsalesiana.ec/ing33ar6r17>
- [18] B. S. Chandar, P. Rajagopalan, and P. Ranganathan, "Short-term aqi forecasts using machine/deep learning models for san francisco, ca," in *2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC)*, 2023, pp. 0402–0411. [Online]. Available: <https://doi.org/10.1109/CCWC57344.2023.10099064>
- [19] J. Wang, P. Du, Y. Hao, X. Ma, T. Niu, and W. Yang, "An innovative hybrid model based on outlier detection and correction algorithm and heuristic intelligent optimization algorithm for daily air quality index forecasting," *Journal of Environmental Management*, vol. 255, p. 109855, 2020. [Online]. Available: <https://doi.org/10.1016/j.jenvman.2019.109855>
- [20] M. Méndez, M. G. Merayo, and M. Núñez, "Machine learning algorithms to forecast air quality: a survey," *Artificial Intelligence Review*, vol. 56, no. 9, pp. 10031–10066, Sep 2023. [Online]. Available: <https://doi.org/10.1007/s10462-023-10424-4>
- [21] C. Ji, C. Zhang, L. Hua, H. Ma, M. S. Nazir, and T. Peng, "A multi-scale evolutionary deep learning model based on ceemdan, improved whale optimization algorithm, regularized extreme learning machine and lstm for aqi prediction," *Environmental Research*, vol. 215, p. 114228, 2022. [Online]. Available: <https://doi.org/10.1016/j.envres.2022.114228>
- [22] R. Yan, J. Liao, J. Yang, W. Sun, M. Nong, and F. Li, "Multi-hour and multi-site air quality index forecasting in beijing using cnn, lstm, cnn-lstm, and spatiotemporal clustering," *Expert Systems with Applications*, vol. 169, p. 114513, 2021. [Online]. Available: <https://doi.org/10.1016/j.eswa.2020.114513>
- [23] E. Hossain, M. A. U. Shariff, M. S. Hossain, and K. Andersson, "A novel deep learning approach to predict air quality index," in *Proceedings of International Conference on Trends in Computational and Cognitive Engineering*, M. S. Kaiser, A. Bandyopadhyay, M. Mahmud, and K. Ray, Eds. Singapore: Springer Singapore, 2021, pp. 367–381. [Online]. Available: https://doi.org/10.1007/978-981-33-4673-4_29

- [24] Q. Wen, T. Zhou, C. Zhang, W. Chen, Z. Ma, J. Yan, and L. Sun, "Transformers in time series: a survey," in *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, ser. IJCAI '23, Macao, P.R.China, 2023. [Online]. Available: <https://doi.org/10.24963/ijcai.2023/759>
- [25] Y. Guo, T. Zhu, Z. Li, and C. Ni, "Auto-modal: Air-quality index forecasting with modal decomposition attention," *Sensors*, vol. 22, no. 18, 2022. [Online]. Available: <https://doi.org/10.3390/s22186953>
- [26] S. Ma, J. He, J. He, Q. Feng, and Y. Bi, "Forecasting air quality index in yan'an using temporal encoded informer," *Expert Systems with Applications*, vol. 255, p. 124868, 2024. [Online]. Available: <https://doi.org/10.1016/j.eswa.2024.124868>
- [27] J. Xie, J. Li, M. Zhu, and Q. Wang, "Multi-step air quality index forecasting based on parallel multi-input transformers," in *Pattern Recognition*, H. Lu, M. Blumenstein, S.-B. Cho, C.-L. Liu, Y. Yagi, and T. Kamiya, Eds. Cham: Springer Nature Switzerland, 2023, pp. 52–63. [Online]. Available: https://doi.org/10.1007/978-3-031-47665-5_5
- [28] U.S. Environmental Protection Agency, *Guideline for Reporting of Daily Air Quality – Air Quality Index (AQI)*. U.S. Environmental Protection Agency Research Triangle Park, North Carolina, 2006. [Online]. Available: <https://upsalesiana.ec/ing33ar6r28>
- [29] H. S. Hota, R. Handa, and A. K. Shrivastava, "Time series data prediction using sliding window based rbf neural network," in *International Journal of Computational Intelligence Research*, vol. 13, no. 5, 2017, pp. 1145–1156. [Online]. Available: <https://upsalesiana.ec/ing33ar6r29>
- [30] D. C. Yadav and S. Pal, "15 - measure the superior functionality of machine intelligence in brain tumor disease prediction," in *Artificial Intelligence-Based Brain-Computer Interface*, V. Bajaj and G. Sinha, Eds. Academic Press, 2022, pp. 353–368. [Online]. Available: <https://doi.org/10.1016/B978-0-323-91197-9.00005-9>
- [31] A. Sherstinsky, "Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network," *Physica D: Nonlinear Phenomena*, vol. 404, p. 132306, 2020. [Online]. Available: <https://doi.org/10.1016/j.physd.2019.132306>
- [32] S. Hesaraki. (2023) Long short-term memory (lstm). Medium. Human Stories ideas. [Online]. Available: <https://upsalesiana.ec/ing33ar6r32>
- [33] Y. Mansar. (2021) How to use transformer networks to build a forecasting model. Medium. Human Stories ideas. [Online]. Available: <https://upsalesiana.ec/ing33ar6r33>
- [34] J. A. Segovia, J. F. Toaquiza, J. R. Llanos, and D. R. Rivas, "Meteorological variables forecasting system using machine learning and open-source software," *Electronics*, vol. 12, no. 4, 2023. [Online]. Available: <https://doi.org/10.3390/electronics12041007>
- [35] M. Hammad Hasan. (2023) Tuning model hyperparameters with random searcha comparison with grid search. Medium. Human Stories ideas. [Online]. Available: <https://upsalesiana.ec/ing33ar6r35>
- [36] B. Cui, M. Liu, S. Li, Z. Jin, Y. Zeng, and X. Lin, "Deep learning methods for atmospheric pm2.5 prediction: A comparative study of transformer and cnn-lstm-attention," *Atmospheric Pollution Research*, vol. 14, no. 9, p. 101833, 2023. [Online]. Available: <https://doi.org/10.1016/j.apr.2023.101833>



COMPARATIVE STUDY OF COOLING STRATEGIES IN A LITHIUM-ION BATTERY MODULE FOR THERMAL RUNAWAY PREVENTION USING CFD

ESTUDIO COMPARATIVO DE ESTRATEGIAS DE ENFRIAMIENTO EN UN MÓDULO DE BATERÍAS DE IONES DE LITIO PARA LA PREVENCIÓN DE LA FUGA TÉRMICA MEDIANTE CFD

Ricardo Carpio-Chilloallo^{1,*} , Edwin Paccha-Herrera¹ 

Received: 16-08-2024, Received after review: 14-11-2024, Accepted: 26-11-2024, Published: 01-01-2025

Abstract

This study investigates the thermal behavior of three lithium-ion battery configurations under thermal runaway conditions, focusing on cooling systems based on air, water, and phase change materials (PCM). The analysis was conducted using sixteen cylindrical 18650 cells, each with a capacity of 2.15 Ah. The battery arrangements include Geometry 1, characterized by an irregular rhomboid shape, and Geometry 2, which adopts an irregular octagonal shape. Numerical simulations were carried out using Computational Fluid Dynamics (CFD) tools in ANSYS Fluent, employing a thermal abuse model rooted in a multidimensional, multiscale approach, and incorporating the empirical Newman-Tiedemann-Gauthier-Kim (NTGK) model. Transient simulations were performed under forced and natural convection scenarios to capture dynamic thermal behavior. The findings reveal that natural air cooling fails to prevent thermal runaway under the studied conditions. In contrast, water and PCM-based cooling systems effectively mitigate thermal runaway risks. Furthermore, forced convection with air and water significantly enhances thermal management and successfully prevents thermal runaway.

Keywords: Thermal abuse, ANSYS, cooling, NTGK, thermal runaway

Resumen

En este estudio se evalúa el comportamiento térmico ante condiciones de fuga térmica de tres arreglos de celdas de iones de litio con distintos sistemas de enfriamiento: aire, agua y material de cambio de fase (PCM). Se utilizaron 16 celdas cilíndricas de tipo 18650 con una capacidad de 2.15 Ah. La geometría 1 tiene una forma de rombo irregular, mientras que las geometrías 2 y 3 tienen una forma de un octágono irregular. Se implementaron simulaciones numéricas CFD empleando el software ANSYS Fluent mediante la aplicación del modelo de abuso térmico basado en un enfoque multidimensional multiescala y el modelo empírico NTGK. Se ejecutaron simulaciones transientes considerando convección forzada y natural. Los resultados muestran que, para las condiciones dadas en el estudio, el enfriamiento natural por aire no previene la fuga térmica, mientras que el agua y PCM sí la previenen, al igual que la convección forzada con aire y agua.

Palabras clave: abuso térmico, ANSYS, enfriamiento, NTGK, fuga térmica

^{1,*}Universidad Nacional de Loja, Loja, Ecuador.
Corresponding author✉: ricardo.carpio@unl.edu.ec.

Suggested citation: Carpio-Chilloallo, R and Paccha-Herrera, E. "Comparative study of cooling strategies in a lithium-ion battery module for thermal runaway prevention using CFD," *Ingenius, Revista de Ciencia y Tecnología*, N.º 33, pp. 76-90, 2025, DOI: <https://doi.org/10.17163/ings.n33.2025.07>.

1. Introduction

Lithium-ion batteries (LIBs) are essential components in a wide array of applications, including portable electronic devices, electric vehicles, smartphones, and medical equipment. Despite their widespread use, these batteries face critical challenges, particularly the need to prevent operation at elevated temperatures. Excessive heat can trigger thermal runaway, an uncontrolled exothermic reaction that poses serious risks, including fires that threaten human safety and compromise device integrity [1].

Lithium-ion batteries (LIBs) have emerged as a reliable energy source, renowned for their high efficiency, energy density, and relatively long lifespan [2]. A typical LIB cell comprises a negative electrode (graphite anode), a positive electrode (lithium metal oxide cathode), a separator, and an electrolyte. The separator is positioned between the electrodes to prevent electrical contact while permitting the passage of ions. The electrolyte facilitates ionic conduction within the cell [3]. The charge and discharge cycles of LIBs are defined by the transfer of lithium ions and electrons, a process known as ionization and oxidation [4].

The operation of LIBs is highly temperature-sensitive, with optimal and safe performance typically achieved within the range of 15 °C to 35 °C [5]. During cycling and usage, a material layer known as the solid electrolyte interface (SEI) forms on the surface of the electrodes. This layer begins to decompose at approximately 100 °C. Additionally, the separator material melts and shrinks at 143 °C. At temperatures exceeding 150 °C, thermal runaway, a hazardous and uncontrolled exothermic reaction, can occur [6–9].

Since their invention, lithium-ion batteries have undergone significant advancements in safety. However, the risk of fire and explosion persists due to the potential for thermal runaway. Thermal runaway occurs when the temperature within the battery rises exponentially, surpassing the system's ability to dissipate heat into the environment effectively. This excessive heat is generated by the exothermic chemical decomposition of materials within the cells [10].

Various cooling systems have been developed to ensure the battery operates within an appropriate temperature range. These systems are typically classified based on the cooling medium used, which may include air, liquid, or phase change material (PCM) [11, 12].

Air cooling systems can be classified into two types: natural air cooling and forced air cooling (using fans) [13, 14].

Liquid cooling systems for batteries can be classified as either direct or indirect. Direct cooling involves direct contact between the coolant and the battery cells. In contrast, indirect cooling entails immersing battery packs in an insulating coolant, such as mineral or silicone oils, which are chemically inert and do not

react with the materials on the cell surfaces [13].

Phase change material (PCM) cooling systems utilize substances that absorb or release substantial amounts of thermal energy during phase transitions, enabling heating or cooling as needed [15].

Numerous studies have investigated air, water, and PCM cooling systems to mitigate the risk of thermal runaway. For instance, Ouyang et al. [16] conducted experiments using one hundred 18650-type lithium-ion batteries and demonstrated that incorporating a 4 mm-thick layer of aerogel effectively reduces the maximum temperature of the batteries.

Zhou et al. [17] investigated a battery thermal management system (BTMS) incorporating heat pipes and phase change liquids to regulate temperature and prevent thermal propagation. Their results demonstrated that the system effectively mitigates overheating and thermal runaway, maintaining the battery temperature below 185 °C and limiting the temperature variation to less than 2.1 °C, even under high discharge rate conditions.

Alghamdi et al. [18] conducted an experimental study on various thermal management systems for LIBs in electric vehicles, focusing on the application of PCM. The findings revealed that using PCM alone results in an average temperature of 85 °C, exceeding the safe operational range. However, the incorporation of a thermoelectric module and aluminum fins significantly reduced the average temperature to 48 °C, enhancing the system's safety.

Wu et al. [19] developed a battery thermal management system (BTMS) for electric and hybrid vehicles, utilizing phase change materials (PCM) in combination with heat pipes. Experimental results indicated that the incorporation of heat pipes significantly improves temperature distribution under high discharge rates, maintaining the maximum temperature below 50 °C and achieving more stable thermal fluctuations during cyclic operation. Furthermore, the study demonstrated that a slight increase in airspeed further reduces cell temperatures, attributed to the PCM's phase transition process.

Liu et al. [20] conducted simulations to analyze thermal runaway behaviors in a pack of twelve prismatic LIBs and evaluated three thermal safety measures to mitigate internal short circuits. The study proposed using paraffin-based PCM to delay thermal runaway propagation between the batteries. Furthermore, the authors demonstrated that incorporating insulation into the PCM could further extend this delay. Additionally, a novel thermal protection method utilizing immersion cooling with boiling fluorinated liquid was introduced.

Li et al. [21] developed a numerical thermal abuse model using a Computational Fluid Dynamics (CFD) approach to investigate thermal propagation in lithium-ion battery packs. The model revealed that the spacing

between cells facilitates accelerated heat transfer; however, this design compromises energy density.

Numerous studies have focused on simulating the electrochemical-thermal coupling of batteries, a complex problem within the multiscale and multiphysics domain. To address this challenge, the multidimensional multiscale (MSMD) method has been employed and implemented in ANSYS Fluent. This approach solves equations at various scales and across multiple domains [22], enabling the simulation of battery packs with electrically connected cells [23]. Furthermore, the MSMD method can model thermal runaway in batteries under thermal abuse and short-circuit conditions [24]. Paccha-Herrera et al. [25] further advanced this approach by integrating the empirical NTGK model, which facilitates the calculation of thermal and electrical properties, thereby enhancing the assessment of the thermal performance of LIBs.

This study aims to analyze strategies for preventing thermal runaway by evaluating various cooling methods and geometric arrangements of cell models. The primary motivation is to mitigate the risk of thermal runaway in lithium-ion batteries when operating outside the prescribed temperature range, thereby reducing the likelihood of fires or explosions.

This study presents proposed solutions for prevent-

ing thermal runaway in lithium-ion batteries through the implementation of advanced cooling systems, using CFD-based analysis and simulation. Various geometric configurations of battery cells and cooling strategies are explored to enhance safety and efficiency. These findings aim to contribute to the development of safer and more reliable lithium-ion battery systems.

2. Materials and methods

This study employs three distinct geometric arrangements of battery cells with different cooling systems. Numerical simulations are conducted using the ANSYS STUDENT 2024 R1 software package, leveraging Computational Fluid Dynamics (CFD) models for detailed analysis.

Figure 1 provides a schematic representation of the methodology employed in this study to evaluate cooling strategies for preventing thermal runaway. The process begins with the selection of the battery cell and cooling systems. Next, geometries and meshing are developed. Subsequently, appropriate CFD models are chosen, and materials, initial and boundary conditions, and the numerical solver are configured. The procedure concludes with a comprehensive analysis of the simulation results.

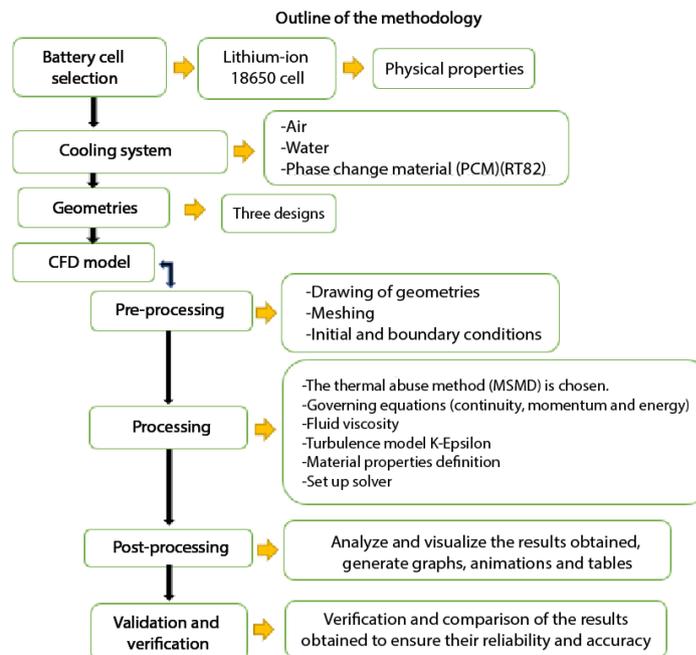


Figure 1. Schematic representation of the CFD study methodology for preventing thermal runaway

2.1. Numerical resolution

The governing equations utilized in the CFD simulations for the air and water domains include the continuity, momentum, and energy equations. The simulations

assume incompressible flow conditions, under which fluid density remains constant. The continuity equation is expressed as Equation (1) [26]:

$$\nabla \cdot \vec{v} = 0 \quad (1)$$

Where ∇ is the divergence operator and \vec{v} is the fluid velocity vector.

Equation (2) represents the momentum conservation equation for an incompressible, viscous fluid:

$$\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \nabla) \vec{v} = \frac{\nabla p}{\rho} + \frac{\mu}{\rho} \nabla^2 \vec{v} \quad (2)$$

Equation (3) illustrates the energy conservation equation for a fluid:

$$\frac{\partial(\rho e)}{\partial t} + \nabla(\rho e \vec{v}) = -p \nabla \cdot \vec{v} + \nabla \cdot (k_f \nabla T) + \Phi \quad (3)$$

Where \vec{v} is the fluid velocity, p is the pressure, ρ is the air density, μ is the dynamic viscosity, e is the internal energy, T is the temperature, k_f is the thermal conductivity of the fluid, Φ is the viscous dissipation, and t is time.

To model the effects of turbulence, the $k - \epsilon$ model was used due to its robustness [27].

In the case of PCM RT82, where a phase change occurs, the continuity equation (Equation (4)), momentum equation (Equation (5)), and energy equation (Equation (6)) are applied [28].

$$\frac{\partial(\rho_f)}{\partial t} + \nabla \cdot (\rho_f \vec{U}) = 0 \quad (4)$$

Where $\frac{\partial \rho_f}{\partial t}$ is the rate of temporal change of density, ρ_f is the fluid phase density, and \vec{U} is the fluid velocity vector (with components \vec{u} , \vec{v} and \vec{w}).

$$\rho_f \frac{\partial \vec{U}}{\partial t} + \rho_f (\vec{U} \cdot \nabla) \vec{U} = -\nabla p + \mu \nabla^2 \vec{U} + \frac{(1 - f_l)^2}{f_l^3 + \delta} A_m \vec{U} + \rho_f g \beta (T_f - T_m) \vec{k} \quad (5)$$

Where β is the consecutive number in the transition region (mushy region), A_m is a parameter for the transition region, δ is a small value introduced to prevent division by zero, and f_l is the liquid fraction.

$$\frac{\partial \rho_f H_f}{\partial t} + \nabla \cdot (\rho_f \vec{u} H_f) = k_f \nabla^2 T_f \quad (6)$$

Where H_f is the fluid enthalpy.

2.2. NTGK model

The thermal runaway phenomenon was modeled using Multiscale Multidomain Modeling (MSMD), based on the empirical NTGK/DCIR model.

The NTGK model, described by Equations (7), (8) and (9), enables the calculation of the thermoelectric properties of the anode, cathode, and active zone domains within a cell or electrically connected battery pack [25]:

$$\frac{\partial(\rho C_p T)}{\partial t} - \nabla \cdot (k_c \nabla T) = \sigma_{\text{pos}} |\nabla \phi_{\text{pos}}|^2 + \sigma_{\text{gen}} |\nabla \phi_{\text{neg}}|^2 + q_{\text{ech}} \quad (7)$$

$$\nabla \cdot (\sigma_{\text{pos}} \nabla \phi_{\text{pos}}) = -j \quad (8)$$

$$\nabla \cdot (\sigma_{\text{neg}} \nabla \phi_{\text{neg}}) = -j \quad (9)$$

Where k_c is the thermal conductivity, σ is the electrical conductivity, ϕ is the electric potential, q_{ech} is the heat transfer rate resulting from thermal exchange with the environment, and the subscripts "pos" and "neg" refer to the positive and negative electrodes, respectively. The volumetric current density j is defined by Equation (10) [29]:

$$j = \frac{C_N}{C_{\text{ref}} \text{Vol}} Y [U - (\phi_{\text{pos}} - \phi_{\text{neg}})] \quad (10)$$

Where Vol is the volume of the active zone, C_{ref} is the battery capacity used to derive the parameters for the functions U and Y , which are defined by Equations (11) and (12), respectively [29]. The corresponding coefficients for these functions are presented in Table 1.

$$U = \left(\sum_{n=0}^5 a_n (D0D)^n \right) - C_2 (T - T_{\text{ref}}) \quad (11)$$

$$Y = \left(\sum_{n=0}^5 b_n (D0D)^n \right) \exp \left[-C_1 \left(\frac{1}{T} - \frac{1}{T_{\text{ref}}} \right) \right] \quad (12)$$

Where C_2 and C_1 are constants for a specific battery.

Table 1. Parámetros para el modelo NTGK [25]

Funciones	
U	Y
$a_0=4.0682$	$b_0=16.5066$
$a_1=-1.2669$	$b_1=-27.0367$
$a_2=-0.9072$	$b_2=237.3297$
$a_3=3.7550$	$b_3=-632.603$
$a_4=-2.3108$	$b_4=725.0825$
$a_5=-0.1701$	$b_5=-309.8760$

The heat transfer rate resulting from thermal exchange with the environment q_{ech} is defined by Equation (13) [25]:

$$q_{\text{ech}} = j \left[U - (\phi_{\text{pos}} - \phi_{\text{neg}}) - T \frac{dU}{dT} \right] \quad (13)$$

Where the first term represents the overpotential heat, and the second term corresponds to the entropic component.

2.3. Thermal abuse model

The thermal abuse model was initially proposed by Harchard et al. [30] and subsequently expanded by Kim et al. [31]. This model is based on the Arrhenius equation and includes a system of equations that describe the reactions occurring during the decomposition of the positive electrode, the negative electrode, and the solid-electrolyte interface (SEI), as outlined below:

$$\frac{dc_{sei}}{dt} = -c_{sei}A_{sei} \cdot \exp\left(-\frac{E_{a,sei}}{RT}\right) \quad (14)$$

$$\frac{dc_{ne}}{dt} = -c_{ne}A_{ne} \cdot \exp\left(-\frac{z_{sei}}{z_{sei,0}}\right) \cdot \exp\left(-\frac{E_{a,ne}}{RT}\right) \quad (15)$$

$$\frac{dz_{sei}}{dt} = c_{ne}A_{ne} \cdot \exp\left(-\frac{z_{sei}}{z_{sei,0}}\right) \cdot \exp\left(-\frac{E_{a,ne}}{RT}\right) \quad (16)$$

$$\frac{d\alpha}{dt} = \alpha(1 - \alpha)A_{pe} \cdot \exp\left(-\frac{E_{a,pe}}{RT}\right) \quad (17)$$

$$\frac{dc_{el}}{dt} = -c_{el}A_{el} \cdot \exp\left(-\frac{E_{a,el}}{RT}\right) \quad (18)$$

Where the subscripts *ne*, *np* and *el* represent the reaction between the negative electrode and the electrolyte, the reaction between the positive electrode and the electrolyte, and the decomposition reaction of the electrolyte, respectively. Additionally, z_{sei} is a dimensionless measure of the SEI layer thickness; $z_{(sei,0)}$ is the reference thickness of the SEI layer; c_{sei} is the fraction of the concentration of metastable lithium-containing species in the SEI layer; c_{ne} is the fraction of the lithium concentration in the negative electrode; α represents the conversion degree of the cathode, and c_{el} is the concentration of the electrolyte. All these variables are dimensionless. R is the universal gas constant.

2.4. Battery Cell Selection

In this study, 18650-type cylindrical lithium-ion cells were used. The properties are detailed in Table 2.

Table 2. Properties of the 18650 lithium-ion cell [26]

Parameters	Units	Value
Diameter	mm	18.4
Height	mm	65
Nominal capacity	Ah	2.15
Nominal voltage	V	3.62
Maximum charge rate	-	1 C
Maximum discharge rate	-	4.65 C
Internal resistance	$m \cdot \Omega$	35
Density	kg/m^3	1852
Heat capacity	$J/kg \cdot K$	1200
Thermal conductivity	-	-
Axial	$W/m \cdot K$	0.2
Radial	$W/m \cdot K$	37.6

Figure 2 illustrates the main components of the 18650 lithium-ion battery used in this study. The battery features a diameter of 18.4 mm, an anode height of 2 mm, a cathode height of 3 mm, and an active zone of 60 mm.

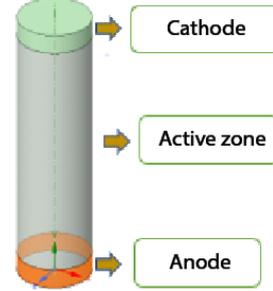


Figure 2. Main components of a lithium-ion battery used in the MSMD approach

2.5. Cooling systems

The cooling systems proposed in this study include air, water, and PCM. For the PCM, RT82 is employed, which consists of pure organic materials capable of storing and releasing significant amounts of heat through a solid-to-liquid or liquid-to-solid phase change process [32]. The properties of the RT82 PCM are detailed in Table 3.

Table 3. Properties of the RT82 PCM [32]

Thermophysical properties	Units	Value
Density (solid)	$kg \cdot m^{-3}$	950
Density (liquid)	$kg \cdot m^{-3}$	770
Specific heat	$J \cdot (kg^{-1} \cdot K^{-1})$	2000
Thermal conductivity	$W \cdot (m^{-1} \cdot K^{-1})$	0.2
Latent heat	J/kg	176000
Dynamic viscosity	$kg/m \cdot s$	0.03499
Temperature (solid)	K	350.15
Temperature (liquid)	K	358.15
Thermal expansion coefficient	1/K	0.001

2.6. Geometric arrangements of the cells

The geometries used in this study were designed using the Ansys Space Claim module within ANSYS Fluent.

Three geometries, each consisting of sixteen 18650 lithium-ion cells, were proposed with varying arrangements, using three different cooling methods: air, water, and PCM RT82. The thermal runaway behavior of the cells was analyzed using the thermal abuse model in ANSYS Fluent, which is based on the Multiscale Multidomain (MSMD) method and incorporates the empirical NTGK model.

The ambient and fluid inlet temperature was set at 25 °C, with a convection coefficient of $7 \frac{W}{m^2K}$ between the casing and the external environment, as recommended by Liu et al. [20]. An inlet velocity of $1.5 \frac{m}{s}$ was applied for both air and water. The convergence criterion was defined as residuals reaching values below 1×10^{-04} . For each simulation case, an acrylic box was used to enclose the batteries. The sixteen cells were connected in series using the MSMD virtual connection in ANSYS Fluent, with a time step size of 1 s. Aluminum was selected as the material to simulate the anode and cathode.

A structured hexahedral mesh was employed to achieve an optimal balance between computational accuracy and efficiency. A mesh independence test was performed using three different element sizes for each geometry, selecting the size that maintained a consistent temperature as the number of elements increased. The number of elements for the various geometries and the average orthogonal quality are presented in Table 4. The values of the orthogonal quality parameter indicate that the mesh quality is excellent.

Table 4. Mesh Characteristics

Geometries	Number of elements	Average orthogonal quality
Geometry 1	269930	0.79
Geometry 2	201794	0.80
Geometry 3	249264	0.80

For the natural convection case, the PRESTO! method was employed for pressure discretization, while a COUPLED scheme was utilized for pressure-velocity coupling.

Geometry 1 features an irregular diamond shape (Figure 3) with a row separation of 48 mm and a column separation of 38 mm. Geometry 2 adopts an irregular octagonal shape (Figure 6) with both row and column separations set at 24 mm. Geometry 3 also exhibits an irregular octagonal shape (Figure 9) but with a row separation of 38 mm and a column separation of 48 mm.

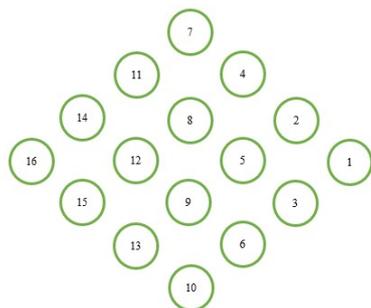


Figure 3. Cell arrangement for Geometry 1

Figures 3, 4 and 5 illustrate the numbered arrangement of the cells, the configuration of the spaces be-

tween the cells and the casing, and the direction of coolant flow for Geometry 1, respectively. Similarly, Figures 6, 7 and 8 depict these features for Geometry 2, while Figures 9, 10 and 11 present the corresponding details for Geometry 3.

This study was conducted using a laptop with a Ryzen 7 processor (7.8 GHz) and 16 GB of RAM.

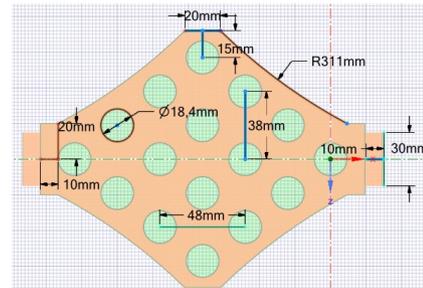


Figure 4. Geometry 1: Dimensions

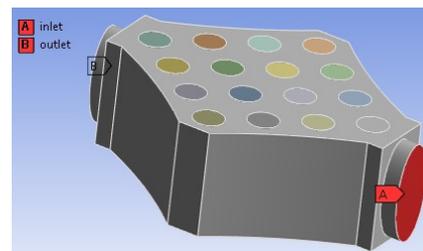


Figure 5. Geometry 1: Coolant flow

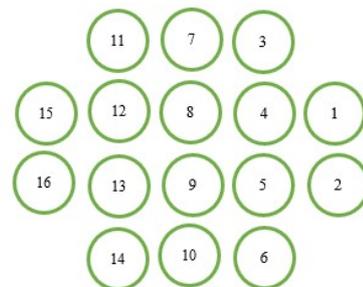


Figure 6. Cell arrangement for Geometry 2

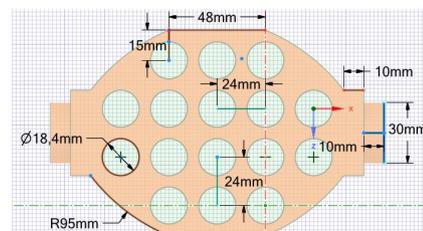


Figure 7. Geometry 2: Dimensions

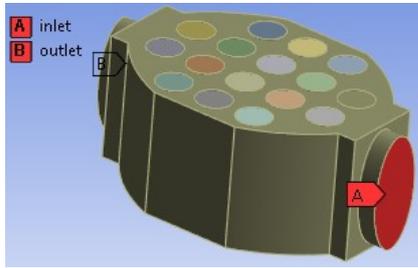


Figure 8. Geometry 2: Coolant flow

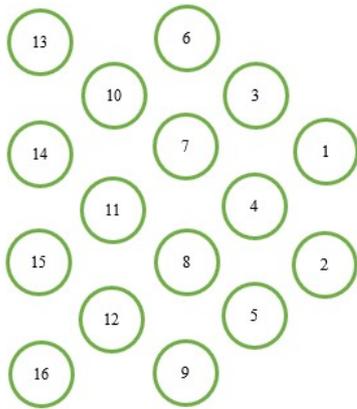


Figure 9. Cell arrangement for Geometry 3

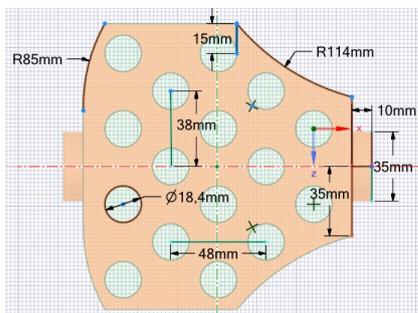


Figure 10. Geometry 3: Dimensions

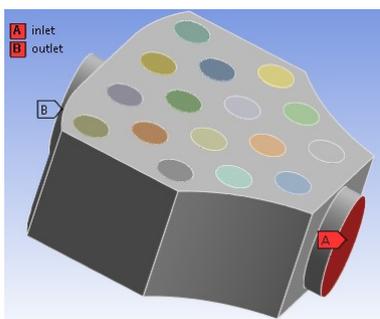


Figure 11. Geometry 3: Coolant flow

3. Results and discussion

Figure 12 presents a representative graph illustrating the cooling process of a cell using air, demonstrating that steady-state conditions are achieved by the end of the process.

Figure 13 illustrates a typical pattern of thermal runaway, where a minor temperature increase after 505 seconds triggers an uncontrolled chain reaction, resulting in an exponential rise in the system's temperature. This phenomenon is particularly critical in lithium-ion batteries, as it can lead to catastrophic failures if not effectively managed.

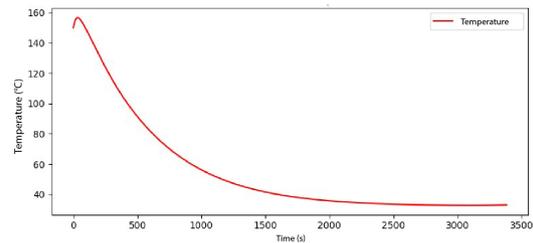


Figure 12. Cooling dynamics of a battery. Air cooling with forced convection, ambient and inlet temperature of 25°C, and inlet velocity of $1.5 \frac{m}{s}$

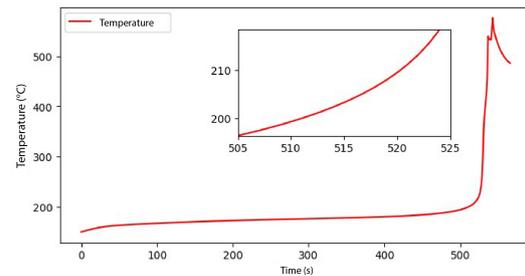


Figure 13. Characteristic temperature profile of thermal runaway in a cell. Air cooling, with an inlet temperature to the module of 47°C and velocity of $1.5 \frac{m}{s}$

3.1. Air cooling

Figure 14 illustrates the temperature profiles of the cells in Geometry 1 under natural convection air cooling. This method fails to prevent thermal runaway, with Cell 16 being the first to exhibit this phenomenon at 258 seconds. Figure 15 depicts the temperature distribution across the cells.

Figure 16 presents the results for Geometry 2 with natural convection air cooling, which fails to prevent thermal runaway. Cells 15 and 16 are the first to exhibit this abnormal condition, occurring after 228 seconds. The temperature distribution across the cells is depicted in Figure 17.

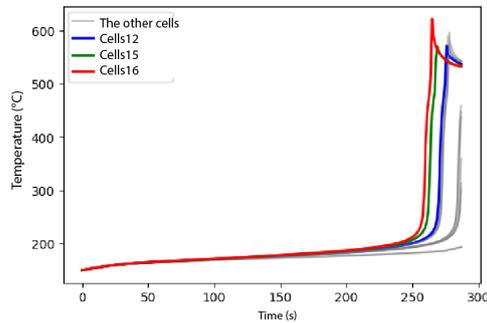


Figure 14. Geometry 1. Air cooling under natural convection

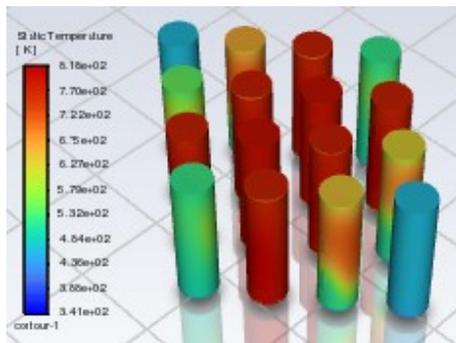


Figure 15. Geometry 1. Cell temperature at 287 s through natural air convection

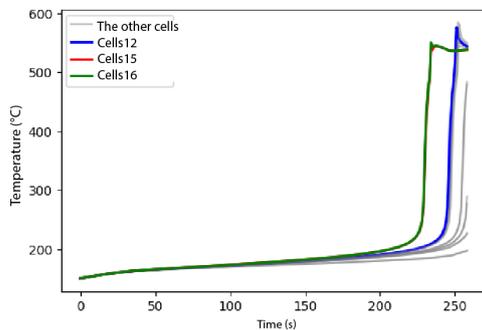


Figure 16. Geometry 2. Air cooling under natural convection

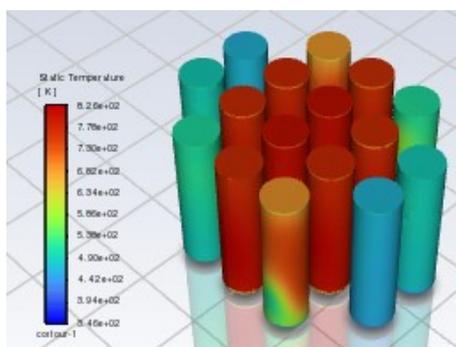


Figure 17. Geometry 2. Cell temperature at 258 s through natural air convection

Figure 18 illustrates the battery temperature profiles for Geometry 3 under natural convection air cooling, with Cell 15 being the first to undergo thermal runaway after 255 seconds. The temperature distribution across the cells is more clearly depicted in Figure 19, highlighting that the central cells experience the most significant temperature increases.

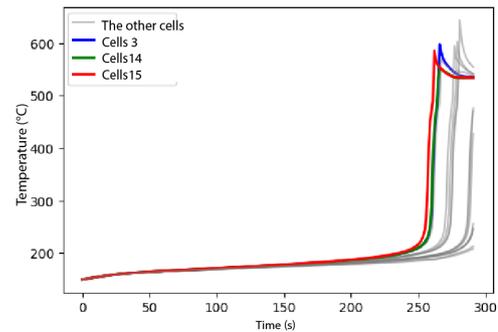


Figure 18. Geometry 3. Air cooling through natural convection

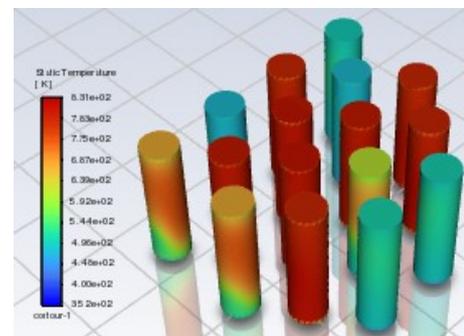


Figure 19. Geometry 3. Cell temperature at 292 s through natural air convection

Figure 20 illustrates the temperature curves for Geometry 1 under forced convection air cooling, which successfully prevents thermal runaway. Cell 5 reaches the highest temperature of 156.98°C within 35 seconds. Figure 21 depicts the temperature distribution across the cells, highlighting insufficient cooling in the cells located at the extremes.

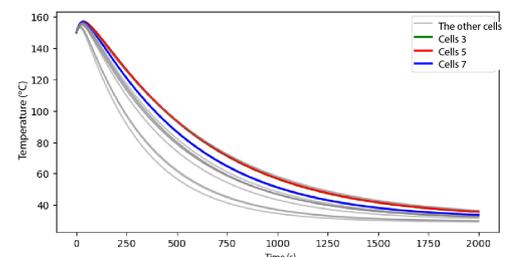


Figure 20. Geometry 1. Forced air convection

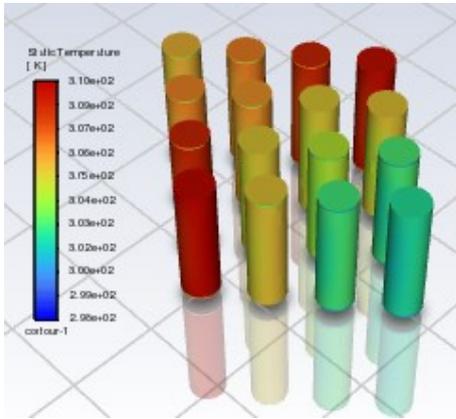


Figure 21. Geometry 1. Cell temperature at 2000 s through forced air convection

Figure 22 presents the temperature evolution of the cells for Geometry 2 under forced convection air cooling, which effectively inhibits thermal runaway. Cell 3 reaches the highest temperature of 157.30°C within 33 seconds. Figure 23 depicts the temperature distribution within the module, indicating that cells located near the air inlet are better cooled compared to those positioned near the air outlet.

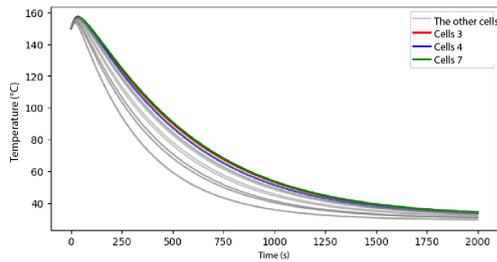


Figure 22. Geometry 2. Forced air convection

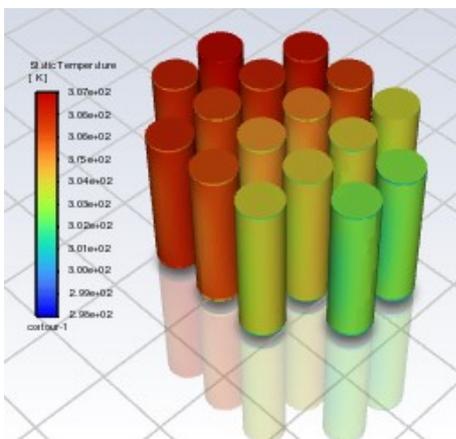


Figure 23. Geometry 2. Cell temperature at 2000 s through forced air convection

Figure 24 illustrates the results for Geometry 3 under forced convection air cooling, which successfully

prevents thermal runaway. However, a greater temperature gradient is observed between neighboring cells compared to Geometry 2. Cell 13 reaches the highest temperature of 159.31°C within 56 seconds. Figure 25 depicts the temperature distribution across the cells, showing that those in the central zone of the module benefit from better cooling.

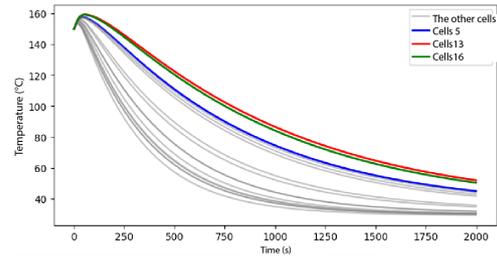


Figure 24. Geometry 3. Forced air convection

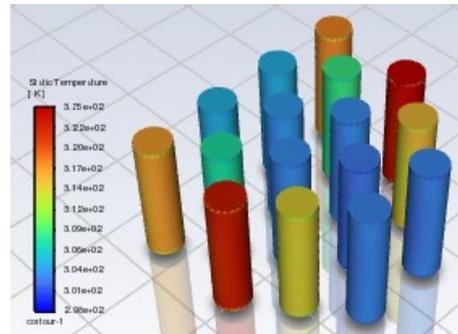


Figure 25. Geometry 3. Cell temperature at 2000 s through forced air convection

3.2. Water cooling

Figure 26 depicts the temperature decay of the cells in Geometry 1, with water cooling via natural convection successfully inhibiting thermal runaway. Cell 16 reaches the highest temperature of 155.90°C within 27 seconds. Figure 27 illustrates the temperature distribution across the cells, demonstrating an almost uniform temperature profile.

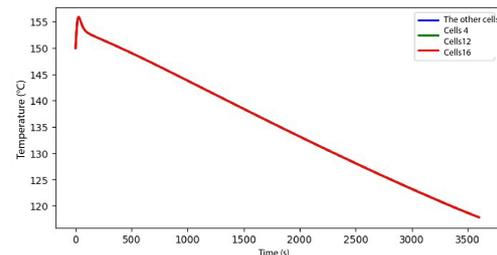


Figure 26. Geometry 1. Water cooling with natural convection

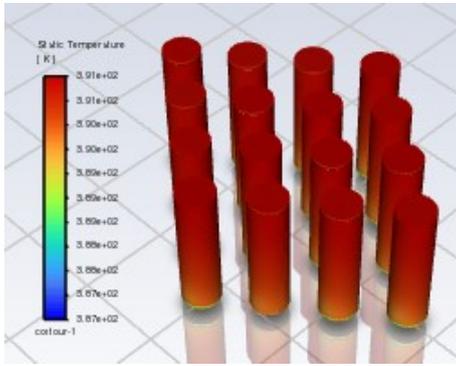


Figure 27. Geometry 1. Cell temperature at 3600 s through natural water convection

Figure 28 illustrates the thermal behavior of the cells in Geometry 2 under water cooling via natural convection, which effectively prevents thermal runaway. Cell 16 reaches the highest temperature of 156.09°C within 29 seconds. Figure 29 shows the temperature distribution across the cells, demonstrating a uniform profile.

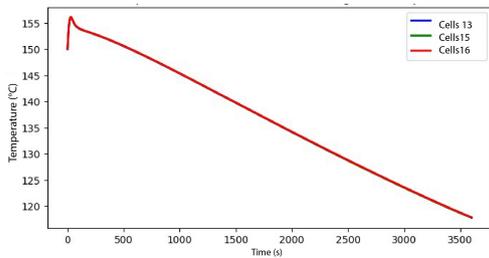


Figure 28. Geometry 2. Water cooling through natural convection

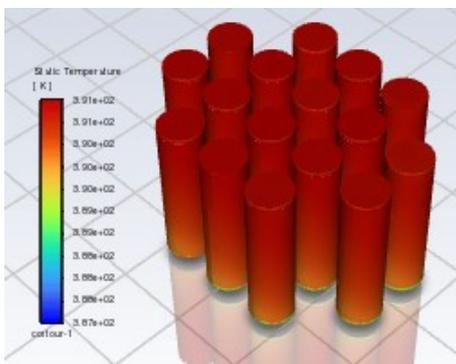


Figure 29. Geometry 2. Cell temperature at 3600 s through natural water convection

Figure 30 presents the results for Geometry 3 with water cooling via natural convection. Cell 7 reaches the highest temperature of 156.19°C within 29 seconds, with no occurrence of thermal runaway. Figure 31 illustrates the uniform temperature distribution across the batteries.

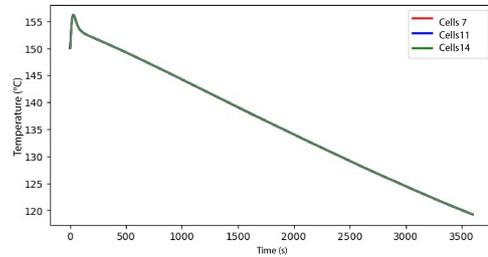


Figure 30. Geometry 3. Water cooling with natural convection

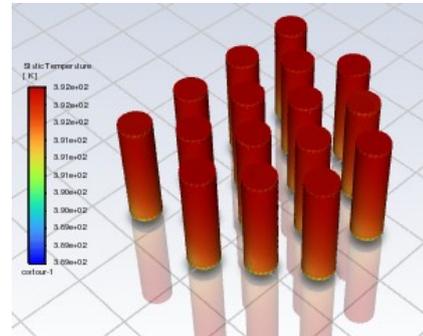


Figure 31. Geometry 3. Cell temperature at 3600 s through natural water convection

Figure 32 illustrates the temperature profiles of selected cells in Geometry 1 with water cooling via forced convection. All cells reach a steady-state condition, effectively inhibiting thermal runaway. Figure 33 depicts a uniform temperature distribution across the cells.

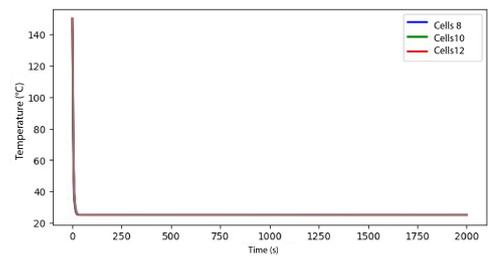


Figure 32. Geometry 1. Water cooling via forced convection

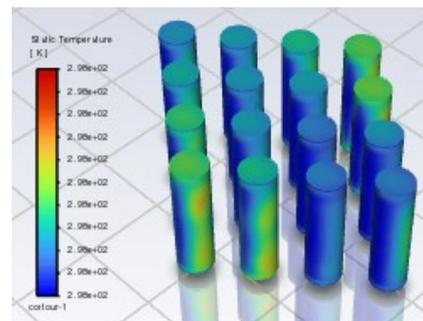


Figure 33. Geometry 1. Cell temperature at 2000 s through forced water convection

Figure 34 illustrates the cell temperature profiles for Geometry 2 with water cooling via forced convection. This cooling method effectively prevents thermal runaway, allowing the system to reach a steady-state condition.

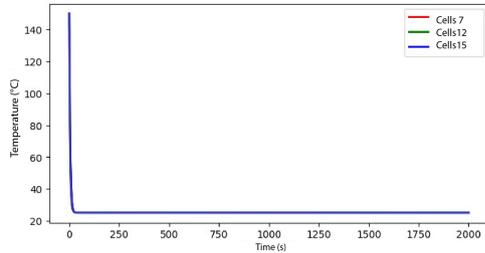


Figure 34. Geometry 2. Water cooling via forced convection

Figure 35 depicts the temperature distribution across the cells, demonstrating a uniform pattern.

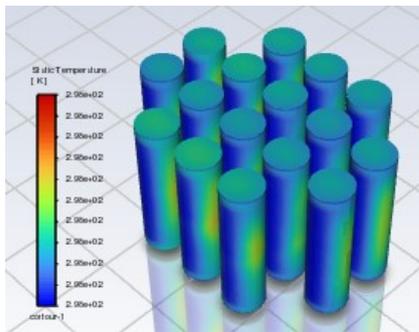


Figure 35. Geometry 2. Cell temperature at 2000 s through forced water convection

Figure 36 presents the results for Geometry 3 with water cooling via forced convection. In this scenario, thermal runaway is effectively inhibited, and a steady-state condition is achieved.

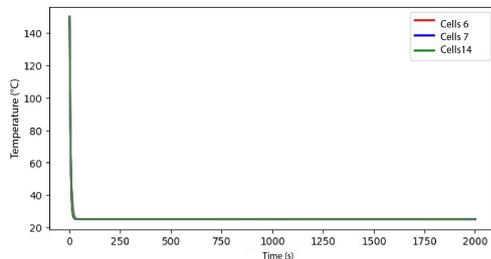


Figure 36. Geometry 3. Water cooling, forced convection

Figure 37 illustrates the temperature distribution across the cells, highlighting negligible temperature gradients.

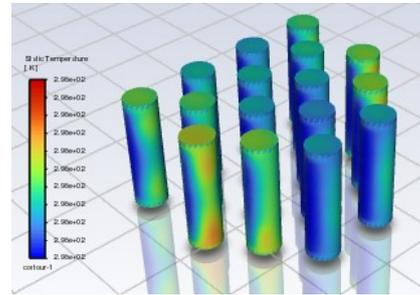


Figure 37. Geometry 2. Cell temperature at 2000 s through forced water convection

3.3. Cooling by PCM (RT82)

Figure 38 illustrates the thermal behavior of Geometry 1 with cooling with PCM (RT82), which successfully prevents thermal runaway. Cell 16 reaches the highest temperature of 161.40°C within 61 seconds, followed by a temperature decay. Figure 39 depicts a uniform temperature distribution across the cells.

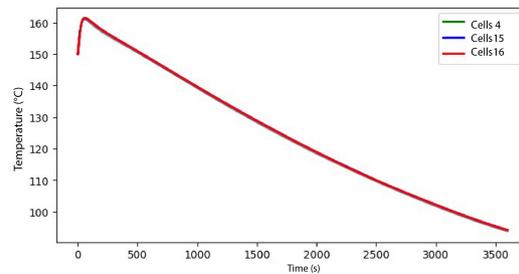


Figure 38. Geometry 1. Cooling by PCM (RT82)

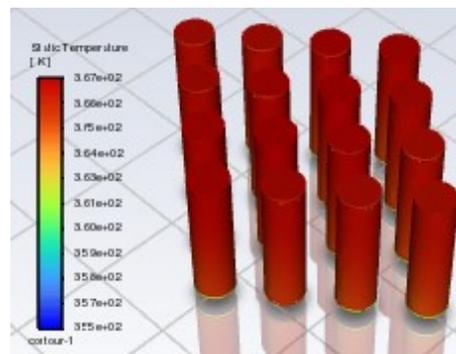


Figure 39. Geometry 1. Cell temperature at 3600 s, cooled by PCM (RT82)

Figure 40 illustrates the cell temperature profiles for Geometry 2 using cooling with PCM (RT82). In this case, thermal runaway is also effectively prevented, with Cells 15 and 16 reaching the highest temperature of 161.71°C within 74 seconds. Figure 41 depicts the temperature distribution across the cells, demonstrating a uniform pattern.

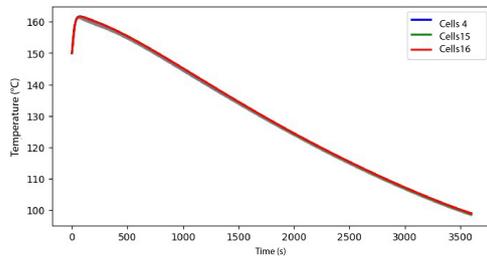


Figure 40. Geometry 2. Cooling by PCM (RT82)

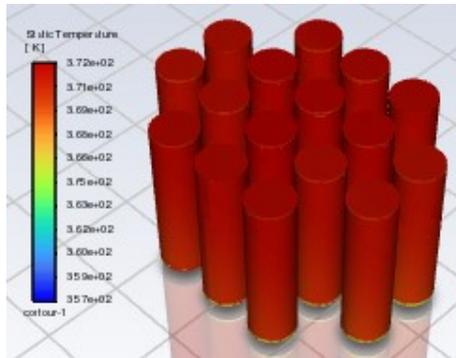


Figure 41. Geometry 2. Cell temperature at 3600 s, cooled by PCM (RT82)

Figure 42 presents the results for Geometry 3 using cooling with PCM. Cell 15 reaches the highest temperature of 161.46°C within 59 seconds, followed by a temperature decay.

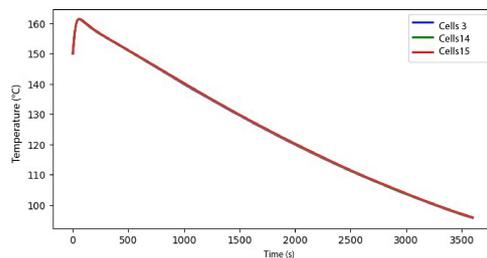


Figure 42. Geometry 3. Cooling by PCM (RT82)

Figure 43 illustrates that the temperature distribution across the cells remains uniform, consistent with the other geometric configurations using PCM.

The phenomenon of thermal runaway plays a critical role in selecting the cooling system and geometric arrangement of battery cells, as its mitigation is highly dependent on these parameters. Consequently, numerous studies have addressed this issue. For instance, Zhou et al. [17] successfully prevented the propagation of thermal runaway, maintaining the temperature below 185°C, with temperatures exceeding 60°C for only 14 seconds. Similarly, Ouyang et al. [16] reduced the maximum temperature of 18650-type batteries from 740.35°C to 55.19°C, effectively preventing thermal runaway. In another study, Alghamdi et al. [18]

employed paraffin-based PCM, achieving an average temperature of 85°C.

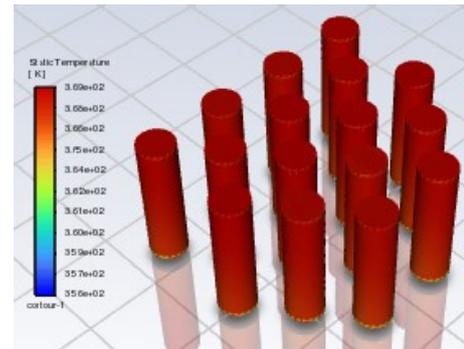


Figure 43. Geometry 3. Cell temperature at 3600 s, cooled by PCM (RT82)

4. Conclusions

A numerical study was conducted to analyze the cooling performance of battery cells within a module utilizing air, water, and phase change material (PCM) as cooling methods. Three distinct geometric configurations of the battery module were proposed to evaluate their effectiveness in inhibiting thermal runaway.

Among the geometries analyzed, none of the three configurations successfully prevented thermal runaway when cooled using natural air convection. However, the phenomenon was effectively inhibited when water and PCM RT82 were employed as the cooling medium. For water cooling, Geometry 1 proved to be the most efficient, with a maximum temperature of 155.90°C. Similarly, when PCM RT82 was used, Geometry 1 also demonstrated the best performance, reaching a maximum temperature of 161.40°C.

Using forced air convection, Geometry 1 emerged as the most efficient configuration, achieving a maximum temperature of 156.98°C. Similarly, with water cooling, Geometry 1 demonstrated superior performance, effectively reducing the temperature from 150°C to 25.06°C.

Based on the results obtained, it can be concluded that Geometry 1 is the most efficient configuration for thermal management in the studied scenarios.

Among the three cooling systems analyzed in this study, water proved to be the most efficient in preventing thermal runaway. This efficiency is attributed to its high heat capacity and thermal conductivity, which enable effective heat distribution and dissipation.

A key limitation of this study is the limited number of cells analyzed. To address this, future work should incorporate models with more diverse geometric configurations and a larger number of cells to improve the robustness of the findings. Additionally, future research should evaluate the effects of indirect contact between

cells and the coolant, as well as explore alternative types of phase change materials (PCM).

This study provides valuable insights into the thermal runaway phenomenon, aiding in the development of more effective cooling systems to prevent it.

References

- [1] R. Gebart, “Thermal runaway criterion for thick polymer composites,” *Composites Part A: Applied Science and Manufacturing*, vol. 182, p. 108187, 2024. [Online]. Available: <https://doi.org/10.1016/j.compositesa.2024.108187>
- [2] A. M. Divakaran, D. Hamilton, K. N. Manjunatha, and M. Minakshi, “Design, development and thermal analysis of reusable li-ion battery module for future mobile and stationary applications,” *Energies*, vol. 13, no. 6, 2020. [Online]. Available: <https://doi.org/10.3390/en13061477>
- [3] G. E. Blomgren, “The development and future of lithium ion batteries,” *Journal of The Electrochemical Society*, vol. 164, no. 1, p. A5019, dec 2016. [Online]. Available: <https://dx.doi.org/10.1149/2.0251701jes>
- [4] A. Carnovale, *Investigation into the Effect of Thermal Management on the Capacity Fade of Lithium-ion Batteries*. Mechanical and Mechatronics Engineering. University of Waterloo, 2016. [Online]. Available: <https://upsalesiana.ec/ing33ar6r4>
- [5] D. Patel, J. B. Robinson, S. Ball, D. J. L. Brett, and P. R. Shearing, “Thermal runaway of a li-ion battery studied by combined arc and multi-length scale x-ray ct,” *Journal of The Electrochemical Society*, vol. 167, no. 9, p. 090511, apr 2020. [Online]. Available: <https://dx.doi.org/10.1149/1945-7111/ab7fb6>
- [6] G. Guo, B. Long, B. Cheng, S. Zhou, P. Xu, and B. Cao, “Three-dimensional thermal finite element modeling of lithium-ion battery in thermal abuse application,” *Journal of Power Sources*, vol. 195, no. 8, pp. 2393–2398, 2010. [Online]. Available: <https://doi.org/10.1016/j.jpowsour.2009.10.090>
- [7] J. Fan, “On the discharge capability and its limiting factors of commercial 18650 li-ion cell at low temperatures,” *Journal of Power Sources*, vol. 117, no. 1, pp. 170–178, 2003. [Online]. Available: [https://doi.org/10.1016/S0378-7753\(03\)00354-9](https://doi.org/10.1016/S0378-7753(03)00354-9)
- [8] K. Amin, J. Zhang, H.-Y. Zhou, R. Lu, M. Zhang, N. Ashraf, C. YueLi, L. Mao, C. F. J. Faul, and Z. Wei, “Surface controlled pseudo-capacitive reactions enabling ultra-fast charging and long-life organic lithium ion batteries,” *Sustainable Energy Fuels*, vol. 4, pp. 4179–4185, 2020. [Online]. Available: <http://dx.doi.org/10.1039/D0SE00610F>
- [9] K. Märker, C. Xu, and C. P. Grey, “Operando nmr of nmc811/graphite lithium-ion batteries: Structure, dynamics, and lithium metal deposition,” *Journal of the American Chemical Society*, vol. 142, no. 41, pp. 17 447–17 456, Oct 2020. [Online]. Available: <https://doi.org/10.1021/jacs.0c06727>
- [10] R. D. McKerracher, J. Guzman-Guemez, R. G. A. Wills, S. M. Sharkh, and D. Kramer, “Advances in prevention of thermal runaway in lithium-ion batteries,” *Advanced Energy and Sustainability Research*, vol. 2, no. 5, p. 2000059, 2021. [Online]. Available: <https://doi.org/10.1002/aesr.202000059>
- [11] Z. Rao and S. Wang, “A review of power battery thermal energy management,” *Renewable and Sustainable Energy Reviews*, vol. 15, no. 9, pp. 4554–4571, 2011. [Online]. Available: <https://doi.org/10.1016/j.rser.2011.07.096>
- [12] H. Behi, D. Karimi, R. Youssef, M. Suresh Patil, J. Van Mierlo, and M. Bercibar, “Comprehensive passive thermal management systems for electric vehicles,” *Energies*, vol. 14, no. 13, 2021. [Online]. Available: <https://doi.org/10.3390/en14133881>
- [13] T. I. C. Buidin and F. Mariasiu, “Battery thermal management systems: Current status and design approach of cooling technologies,” *Energies*, vol. 14, no. 16, 2021. [Online]. Available: <https://doi.org/10.3390/en14164879>
- [14] J. Du, Y. Sun, Y. Huang, and X. Wu, “Analysis of influencing factors of thermal management system for lifepo4 lithium battery under high power charging,” *World Electric Vehicle Journal*, vol. 11, no. 2, 2020. [Online]. Available: <https://doi.org/10.3390/wevj11020044>
- [15] E. Grimonia, M. R. C. Andhika, M. F. N. Aulady, R. V. C. Rubi, and N. L. Hamidah, “Thermal management system using phase change material for lithium-ion battery,” *Journal of Physics: Conference Series*, vol. 2117, no. 1, p. 012005, nov 2021. [Online]. Available: <https://dx.doi.org/10.1088/1742-6596/2117/1/012005>
- [16] T. Ouyang, B. Liu, C. Wang, J. Ye, and S. Liu, “Novel hybrid thermal management system for preventing li-ion battery thermal runaway using nanofluids cooling,” *International Journal of Heat and Mass Transfer*, vol. 201, p. 123652, 2023. [Online]. Available: <https://doi.org/10.1016/j.ijheatmasstransfer.2022.123652>

- [17] H. Zhou, C. Dai, Y. Liu, X. Fu, and Y. Du, "Experimental investigation of battery thermal management and safety with heat pipe and immersion phase change liquid," *Journal of Power Sources*, vol. 473, p. 228545, 2020. [Online]. Available: <https://doi.org/10.1016/j.jpowsour.2020.228545>
- [18] H. Alghamdi, M. F. M. Rosdi, A. Mukhtar, A. S. H. M. Yasir, and A. Alviz-Meza, "Controlling thermal runaway by simultaneous use of thermoelectric module and phase change material in the lithium-ion batteries of electric vehicles," *Case Studies in Thermal Engineering*, vol. 52, p. 103697, 2023. [Online]. Available: <https://doi.org/10.1016/j.csite.2023.103697>
- [19] W. Wu, X. Yang, G. Zhang, K. Chen, and S. Wang, "Experimental investigation on the thermal performance of heat pipe-assisted phase change material based battery thermal management system," *Energy Conversion and Management*, vol. 138, pp. 486–492, 2017. [Online]. Available: <https://doi.org/10.1016/j.enconman.2017.02.022>
- [20] X. Liu, Z. Zhou, W.-T. Wu, L. Wei, W. Wu, Y. Li, L. Gao, Y. Li, and Y. Song, "Modelling for the mitigation of lithium ion battery thermal runaway propagation by using phase change material or liquid immersion cooling," *Case Studies in Thermal Engineering*, vol. 52, p. 103749, 2023. [Online]. Available: <https://doi.org/10.1016/j.csite.2023.103749>
- [21] Y. Li, G. Liu, and Z. Li, "Numerical modeling of thermal runaway in high-energy lithium-ion battery packs induced by multipoint heating," *Case Studies in Thermal Engineering*, vol. 38, p. 102335, 2022. [Online]. Available: <https://doi.org/10.1016/j.csite.2022.102335>
- [22] R. Ni, D. Zhang, R. Wang, Z. Xie, and Y. Wang, "Prevention and suppression effects of phase change material on thermal runaway in batteries," *Case Studies in Thermal Engineering*, vol. 48, p. 103160, 2023. [Online]. Available: <https://doi.org/10.1016/j.csite.2023.103160>
- [23] Y. Liu, Y. G. Liao, and M.-C. Lai, "Transient temperature distributions on lithium-ion polymer sli battery," *Vehicles*, vol. 1, no. 1, pp. 127–137, 2019. [Online]. Available: <https://doi.org/10.3390/vehicles1010008>
- [24] Q. Li, C. Yang, S. Santhanagopalan, K. Smith, J. Lamb, L. A. Steele, and L. Torres-Castro, "Numerical investigation of thermal runaway mitigation through a passive thermal management system," *Journal of Power Sources*, vol. 429, pp. 80–88, 2019. [Online]. Available: <https://doi.org/10.1016/j.jpowsour.2019.04.091>
- [25] E. Paccha-Herrera, W. R. Calderón-Muñoz, M. Orchard, F. Jaramillo, and K. Medjaher, "Thermal modeling approaches for a licoo2 lithium-ion battery—a comparative study with experimental validation," *Batteries*, vol. 6, no. 3, 2020. [Online]. Available: <https://doi.org/10.3390/batteries6030040>
- [26] E. Paccha-Herrera, A. Medina-Sisalima, D. Tapiaperalta, B. Solórzano-Castillo, J. Gómez-Peña, and I. Coronel-Villavicencio, "A numerical study on the influence of geometric cell configurations over the thermal performance in a lithium-ion battery module," *IOP Conference Series: Earth and Environmental Science*, vol. 1370, no. 1, p. 012002, jul 2024. [Online]. Available: <https://dx.doi.org/10.1088/1755-1315/1370/1/012002>
- [27] B. Launder and D. Spalding, "The numerical computation of turbulent flows," *Computer Methods in Applied Mechanics and Engineering*, vol. 3, no. 2, pp. 269–289, 1974. [Online]. Available: [https://doi.org/10.1016/0045-7825\(74\)90029-2](https://doi.org/10.1016/0045-7825(74)90029-2)
- [28] H. Liu, C. Jin, H. Li, and Y. Ji, "A numerical study of pcm battery thermal management performance enhancement with fin structures," *Energy Reports*, vol. 9, pp. 1793–1802, 2023, selected papers from 2022 International Conference on Frontiers of Energy and Environment Engineering. [Online]. Available: <https://doi.org/10.1016/j.egyrs.2023.04.214>
- [29] K. H. Kwon, C. B. Shin, T. H. Kang, and C.-S. Kim, "A two-dimensional modeling of a lithium-polymer battery," *Journal of Power Sources*, vol. 163, no. 1, pp. 151–157, 2006, special issue including selected papers presented at the Second International Conference on Polymer Batteries and Fuel Cells together with regular papers. [Online]. Available: <https://doi.org/10.1016/j.jpowsour.2006.03.012>
- [30] T. D. Hatchard, D. D. MacNeil, A. Basu, and J. R. Dahn, "Thermal model of cylindrical and prismatic lithium-ion cells," *Journal of The Electrochemical Society*, vol. 148, no. 7, p. A755, jun 2001. [Online]. Available: <https://dx.doi.org/10.1149/1.1377592>
- [31] G.-H. Kim, A. Pesaran, and R. Spotnitz, "A three-dimensional thermal abuse model for lithium-ion cells," *Journal of Power Sources*, vol. 170, no. 2, pp. 476–489, 2007. [Online]. Available: <https://doi.org/10.1016/j.jpowsour.2007.04.018>
- [32] Y. Han, Y. Yang, T. Mallick, and C. Wen, "Nanoparticles to enhance melting performance of phase change materials

for thermal energy storage,” *Nanomaterials*,
vol. 12, no. 11, 2022. [Online]. Available:

<https://doi.org/10.3390/nano12111864>



CONVOLUTIONAL NEURAL NETWORKS FOR DIABETIC RETINOPATHY DETECTION

REDES NEURONALES CONVOLUCIONALES PARA DETECCIÓN DE RETINOPATÍA DIABÉTICA

Darwin Patiño-Pérez^{1,*} , Luis Armijos-Valarezo¹ ,
 Luis Chóez-Acosta¹ , Freddy Burgos-Robalino¹ 

Received: 15-05-2024, Received after review: 13-11-2024, Accepted: 29-11-2024, Published: 01-01-2025

Abstract

The early detection of diabetic retinopathy remains a critical challenge in medical diagnostics, with deep learning techniques in artificial intelligence offering promising solutions for identifying pathological patterns in retinal images. This study evaluates and compares the performance of three convolutional neural network (CNN) architectures ResNet-18, ResNet-50, and a custom, non-pretrained CNN using a dataset of retinal images classified into five categories. The findings reveal significant differences in the models' ability to learn and generalize. The non-pretrained CNN consistently outperformed the pretrained ResNet-18 and ResNet-50 models, achieving an accuracy of 91 % and demonstrating notable classification stability. In contrast, ResNet-18 suffered severe performance degradation, with accuracy dropping from 70 % to 26 %, while ResNet-50 required extensive tuning to improve its outcomes. The non-pretrained CNN excelled in handling class imbalances and capturing complex diagnostic patterns, emphasizing the potential of tailored architectures for medical imaging tasks. These results underscore the importance of designing domain-specific architectures, demonstrating that model complexity does not necessarily guarantee better performance. Particularly in scenarios with limited datasets, well-designed custom models can surpass pre-trained architectures in diagnostic imaging applications.

Keywords: diabetic retinopathy, blindness, detection, artificial intelligence, convolutional neural networks, image analysis

Resumen

La detección temprana de la retinopatía diabética representa un desafío crítico en el diagnóstico médico, donde el aprendizaje profundo dentro del campo de la inteligencia artificial emerge como una herramienta prometedora para optimizar la identificación de patrones patológicos en imágenes retinales. Este estudio evaluó comparativamente tres arquitecturas de redes neuronales convolucionales ResNet-18, ResNet-50 y una CNN personalizada o no-preentrenada para clasificar imágenes de retinopatía diabética en un conjunto de datos de imágenes agrupadas en cinco categorías, revelando diferencias significativas en su capacidad para aprender y generalizar. Los resultados demostraron que la arquitectura de red neuronal convolucional no-preentrenada superó consistentemente a los modelos preentrenados basados en ResNet-18 y ResNet-50, alcanzando una precisión del 91 % y una notable estabilidad en la clasificación. Mientras ResNet-18 mostró limitaciones severas, degradándose de un 70 % a un 26 % de precisión, y ResNet-50 requirió ajustes para mejorar su rendimiento, la CNN no preentrenada exhibió una capacidad sobresaliente para manejar el desbalance de clases y capturar patrones diagnósticos complejos. El estudio subraya la importancia de diseñar arquitecturas específicamente adaptadas a problemas médicos, destacando que la complejidad no garantiza necesariamente un mejor desempeño, y que un diseño cuidadoso puede superar modelos preentrenados en tareas de diagnóstico por imagen cuando la cantidad de datos con que se cuenta es limitada.

Palabras clave: retinopatía diabética, ceguera, detección, inteligencia artificial, redes neuronales convolucionales, imágenes oculares.

^{1,*}Universidad de Guayaquil, Guayaquil, Ecuador. Corresponding author ✉: darwin.patinop@ug.edu.ec.

Suggested citation: Patiño-Pérez, D., Armijos-Valarezo, L., Chóez-Acosta, L. and Burgos-Robalino, F. "Convolutional neural networks for diabetic retinopathy detection," *Ingenius, Revista de Ciencia y Tecnología*, N.º 33, pp. 91-101, 2025, DOI: <https://doi.org/10.17163/ings.n33.2025.08>.

1. Introduction

The retina, located at the back of the eye, is a vital layer of light-sensitive cells essential for vision. Unfortunately, it is susceptible to various diseases, among which diabetic retinopathy (DR) stands out as one of the most common and serious conditions (Figure 1). DR is an ocular complication of diabetes, characterized by damage to the blood vessels in the retina [1]. This vascular damage can lead to several pathological issues, including:

- **Obstruction of blood flow.** Blocked blood vessels hinder sufficient blood supply to the retina, potentially resulting in the death of retinal cells and subsequent vision loss.
- **Blood leakage.** Damaged blood vessels may leak blood and other fluids into the retina, causing swelling and blurred vision.
- **Growth of abnormal blood vessels.** As a response to oxygen deprivation, the retina can develop new abnormal blood vessels which may be fragile and prone to bleeding.

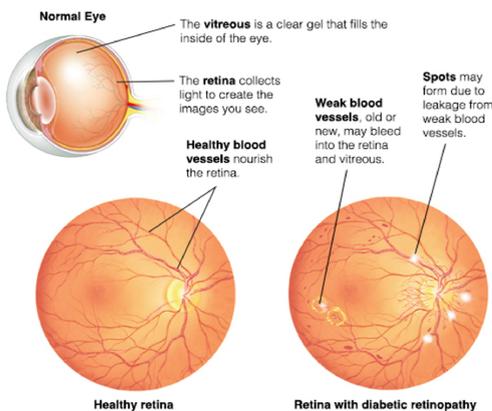


Figure 1. Diabetic retinopathy

Diabetic retinopathy (DR) is more prevalent among individuals with type 1 and type 2 diabetes, particularly those who fail to maintain adequate blood sugar control [2]. Additional risk factors include hypertension, hypercholesterolemia, smoking, overweight or obesity, and pregnancy. In its early stages, DR often presents without noticeable symptoms.

However, as the disease progresses, symptoms may manifest, including blurred vision, dark spots or floaters, difficulty seeing at night, distorted vision, and even vision loss. Early detection and timely treatment are essential to prevent permanent vision impairment. Therefore, regular eye examinations are strongly recommended for individuals with diabetes, particularly those with additional risk factors, to facilitate early intervention and management.

Computer science focuses on designing and developing systems and algorithms capable of performing tasks that typically require human intelligence, such as learning, perception, reasoning, and problem-solving. This discipline forms the foundation of artificial intelligence (AI) [3]. AI integrates techniques from computer science, statistics, logic, and mathematics to create systems that can autonomously learn from data and enhance their performance in real time.

Artificial intelligence (AI) has emerged as a promising tool for DR detection. Machine Learning algorithms can analyze retinal images and identify subtle patterns indicative of the disease. This technology holds significant potential to enhance the accuracy and efficiency of DR diagnosis, enabling earlier detection and facilitating timely interventions.

DR is a severe complication of diabetes that can lead to vision loss if left untreated. Early detection and timely intervention are critical to preventing disease progression and mitigating its impact. Artificial intelligence (AI) has emerged as a promising tool to enhance the detection of DR, offering more precise and efficient diagnostic capabilities and contributing to the preservation of visual health in individuals with diabetes.

Predictive models, which provide forecasts for dichotomous outcomes (distinct yet complementary results), are widely utilized in medical applications. Figure 2 illustrates an evaluation of the most relevant models employed in this domain [4]. Deep learning, a prominent field within artificial intelligence, enables machines or computers to learn and analyze data in a manner akin to human intelligence [5]. This study examines the behavior of various deep learning-based models, highlighting their capability to leverage multiple processing layers to facilitate learning from data representations at multiple levels of abstraction [6].

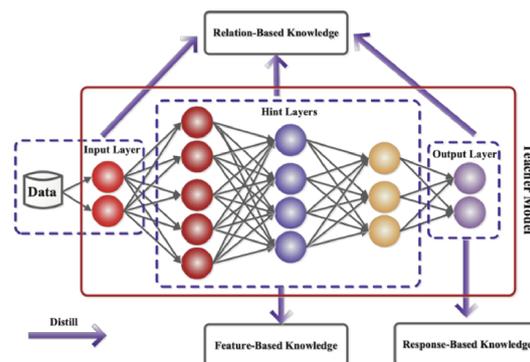


Figure 2. Deep learning model

Numerous pre-trained ResNet implementations are available across various machine learning frameworks, including TensorFlow, PyTorch, Keras, and MXNet. Each framework offers its own variants and specific

optimizations, making the selection of an appropriate pre-trained ResNet model crucial for addressing a particular task. Key factors to consider when selecting a model include the size and complexity of the dataset, the nature of the task (e.g., classification, object detection, or segmentation), and the computational resources available.

The variety of pre-trained ResNet architectures available is extensive, offering a range of options tailored to different tasks and requirements. Selecting the appropriate pre-trained ResNet model depends on the specific objectives of the project and the characteristics of the problem to be addressed. For this study, which focuses on disease recognition through the analysis of ocular images using supervised learning in a classification framework, ResNet models were chosen due to their demonstrated acceptable performance in prior studies involving other types of images. This research evaluates the performance of artificial neural network models, specifically pre-trained convolutional neural networks (CNNs) such as ResNet-18 and ResNet-50, as depicted in Figure 3.

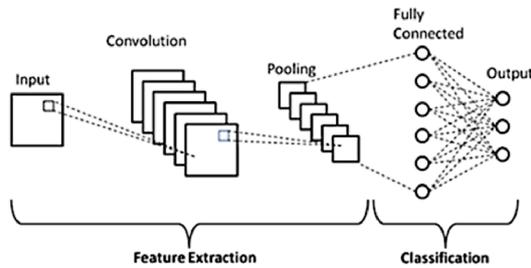


Figure 3. CNN model

2. Materials and methods

2.1. Methodology

Deep learning (DL), a subfield of Machine Learning (ML), represents more than a mere analysis technique (Figure 4). It is a comprehensive methodology that encompasses the entire data science pipeline, including data collection, preparation, exploration, modeling, and evaluation. This approach enables the identification of patterns, the generation of predictions, and informed decision-making. Unlike traditional statistical methods, which rely on predefined rules and static models, ML employs algorithms that learn directly from the data. These algorithms adapt to the data’s complexity and evolve over time, improving performance as they are exposed to larger and more diverse datasets. This adaptability is particularly evident in models based on artificial neural networks, which comprise numerous interconnected neurons organized into layers. These networks follow a hierarchical structure, as depicted in Figure 5.

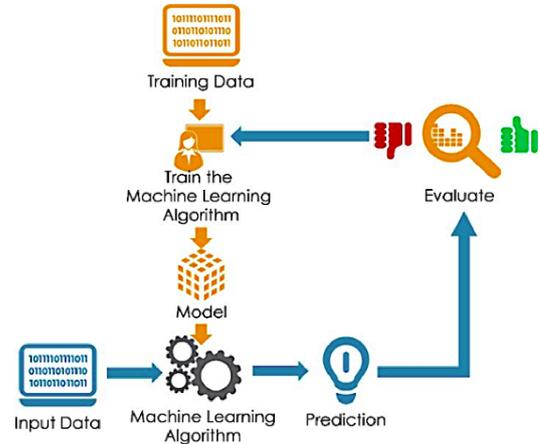


Figure 4. Analysis technique - ML

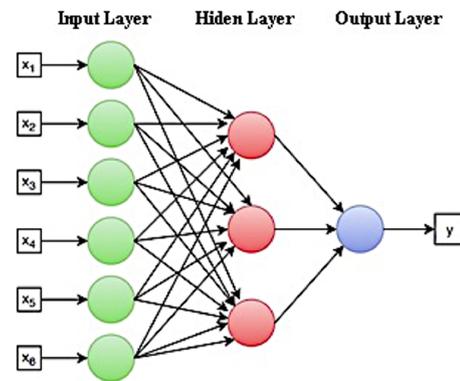


Figure 5. Analysis technique - ML

2.2. The data set

The study dataset utilized in this research comprises 3,662 retinal images sourced from the Kaggle AP-TOS 2019 Blindness Detection (BD) online community. These images are classified based on the severity of diabetic retinopathy, categorized as no diabetic retinopathy, mild, or severe diabetic retinopathy.

Table 1 provides an overview of the dataset, which consists of 3,662 medical images sourced from the Kaggle online community.

Table 1. Medical image segmentation

Type	Name	Number of images	Percentage
0	No DR	1805	49.29%
1	Mild	370	10.10%
2	Moderate	999	27.28%
3	Severe	193	5.27%
4	Proliferative	295	8.06%
Total		3662	100%

2.3. Treatment and adjustment of images

Due to variations in acquisition conditions and equipment, many images in the dataset display differences in retinal alignment and quality. To address these inconsistencies and enable the models to learn network properties more efficiently, an image-processing method was implemented using the OpenCV library in Python.

The preprocessing steps included Gaussian blurring and circular cropping. A contour was drawn around each image, followed by the application of a Gaussian filter. This process reduces high-frequency components, enhancing the clarity of key features in each image and improving their suitability for analysis.

2.4. Description of the variables

Figure 6 illustrates the different levels of diabetic retinopathy (DR), which are categorized as follows:

- **Level 0 (No DR).** This level indicates a non-pathological state, meaning the absence of diabetic retinopathy.
- **Level 1 (Mild).** This stage is characterized by mild non-pathological diabetic retinopathy, where microaneurysms (red spots) are present. These microaneurysms are the source of hard exudates, which appear as high-contrast yellow spots.
- **Level 2 (Moderate).** At this stage, blood vessel distortion and swelling may occur, potentially compromising their ability to transport blood effectively.
- **Level 3 (Severe).** This stage is marked by significant blockage of blood vessels, leading to impaired blood supply to the retina.
- **Level 4 (Proliferative).** This is the most advanced stage, characterized by the secretion of growth factors by the retina, stimulating the proliferation of new blood vessels. These abnormal vessels grow within the retina and extend into the vitreous gel, filling the eye.

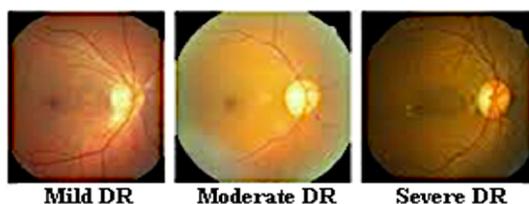


Figure 6. Types of retinopathies

Each stage of diabetic retinopathy has distinct characteristics and properties. However, during analysis,

clinicians may overlook certain details, which could increase the likelihood of an incorrect diagnosis.

2.5. Analysis of data

Initially, the images were downloaded and uploaded to Google Drive. Subsequently, they were organized into directories corresponding to each level of diabetic retinopathy, ensuring accurate differentiation. The critical factor for verifying the correctness of the results lies in the data associated with each retinal image, which is securely stored in the cloud.

The dataset stored in the cloud was used to enable algorithms in Google Colab to access the necessary information for training. The correctness of the results is determined by comparing the output of the classifier model or algorithm with the information available in the cloud. If the results match, it can be concluded that the classification is accurate; otherwise, the result is deemed incorrect. Once the training of each algorithm is completed and the results are obtained, a comparative analysis is conducted to evaluate their performance and identify the most efficient algorithm for solving the proposed problem.

Classification problems in machine learning are broadly divided into two main categories: binary problems and multi-class problems. The key distinction lies in the number of classes the model is required to identify within the data. In the case of binary problems, the model distinguishes between only two classes. These problems are characterized by simplicity, as binary models are generally easier to train and interpret due to the limited number of classes involved. Understanding the model's decisions is also more straightforward, as there are only two possible outcomes. In contrast, multi-class classification involves distinguishing between more than two classes, increasing the complexity of the task. These problems are more challenging to train and interpret due to the larger number of classes and the intricate relationships between them. Convolutional neural networks (CNNs) are particularly well-suited for solving multi-class problems, but interpreting the decisions of such models can be more difficult, given the wider range of potential outcomes.

2.6. Validation metrics

Confusion matrix

The confusion matrix plays a crucial role in identifying errors, enabling both descriptive and analytical evaluations of classification models. It displays the various correct and incorrect assignments made by the model [7]. Using the values provided by the confusion matrix, key evaluation metrics can be calculated to assess the model's performance, as illustrated in Figure 7.

		Predicted Class		
		Positive	Negative	
Actual Class	Positive	True Positive (TP)	False Negative (FN) Type II Error	Sensitivity $\frac{TP}{(TP + FN)}$
	Negative	False Positive (FP) Type I Error	True Negative (TN)	Specificity $\frac{TN}{(TN + FP)}$
		Precision $\frac{TP}{(TP + FP)}$	Negative Predictive Value $\frac{TN}{(TN + FN)}$	Accuracy $\frac{TP + TN}{(TP + TN + FP + FN)}$

Figure 7. Confusion matrix

The confusion matrix is an essential tool for validating neural networks, particularly in classification tasks. It offers detailed insights into the model’s performance by quantifying the number of correct and incorrect predictions for each class.

Among the metrics derived from the confusion matrix and commonly applied to classification tasks in convolutional neural networks (CNNs) are accuracy and loss. These performance metrics are widely used to evaluate image classification models, both in pre-trained convolutional neural networks (pre-trained CNNs) and models developed using scikit-learn (sklearn). However, it is crucial to understand their limitations and to use them in conjunction with other metrics for a more comprehensive evaluation of the model’s performance.

Accuracy

It represents the proportion of correct predictions made by the model, calculated as the number of correct predictions divided by the total number of predictions. It is an intuitive and straightforward metric to interpret; a high accuracy value indicates that the model is generally making accurate predictions. Accuracy is also a useful metric for quickly and easily comparing different models. However, accuracy is sensitive to the distribution of classes. If one class dominates the dataset, the model may achieve high accuracy by predominantly predicting the majority class, even if its performance on other classes is poor. This limitation makes accuracy less reliable in the presence of class imbalance.

Loss

It represents the average error of the model’s predictions and is calculated as the sum of the individual errors for each prediction. It provides insight into the magnitude of the error, where a lower loss value indicates that the model is making predictions with smaller overall errors. Loss plays a crucial role in optimizing the model. During training, it is used to adjust the weights of the neural network to minimize error and

improve performance. The interpretation and scale of the loss depend on the specific loss function used, as different loss functions may have varying meanings and scales. However, loss can be influenced by class imbalance, which should be carefully considered when evaluating the model’s performance.

For pre-trained convolutional neural networks (CNNs), the effectiveness of accuracy and loss metrics depends on the quality of the pre-trained model and its suitability for the specific classification task. Careful selection of the pre-trained network, along with appropriate hyperparameter tuning, is essential to optimize performance and ensure accurate evaluation. In models developed using sklearn, accuracy and loss metrics are directly applicable to classification tasks. However, it is crucial to account for the specific characteristics of the model and the classification problem when selecting appropriate metrics and evaluation techniques.

The effectiveness and reliability of accuracy and loss metrics depend on several factors, including the complexity of the problem, the quality of the data, the model architecture, and the additional metrics employed. It is essential to understand the limitations of these metrics and to use them responsibly in conjunction with other evaluation methods to ensure a comprehensive and robust assessment of image classification models.

2.7. Deep learning models utilized

2.7.1. Pre-trained models

The pre-trained neural networks employed in this study are based on the residual network (ResNet) architecture, which addresses the problem of gradient degradation by incorporating residual blocks. A residual block serves as a fundamental building unit in ResNets and consists of two paths within the network:

1. **Main path:** This path includes the convolutional or fully connected layers typical of a deep neural network.
2. **Direct path:** This is a direct connection that bypasses the layers in the main path, adding its output directly to the output of the main path.

This dual-path structure enables information to propagate through the network without being distorted by the transformations applied in the main path. Consequently, it simplifies the learning process and facilitates the training of much deeper neural networks compared to traditional architectures. For this study, two variants of ResNet were utilized: ResNet-18 and ResNet-50.

ResNet-18

It is an 18-layer deep convolutional neural network, as illustrated in Figure 8. Due to its relatively shallow architecture, ResNet-18 can effectively retain low-scale features, making it appropriate for serving as a feature extractor (encoder). The ResNet-18 architecture comprises 16 convolutional layers, 2 down sampling layers, and several fully connected layers [8].

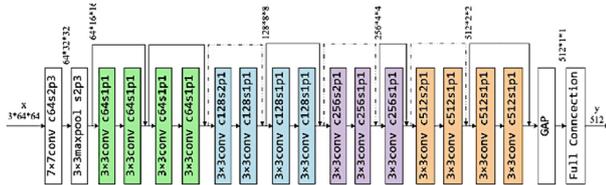


Figure 8. ResNet-18 model

ResNet-50

It is a convolutional artificial neural network with a depth of 50 layers, as depicted in Figure 9. It can utilize a pre-trained version trained on over one million images from the ImageNet database [9]. The ResNet-50 architecture consists of 48 convolutional layers, one MaxPooling layer, and one average pooling layer. It requires approximately 3.8×10^9 floating-point operations.

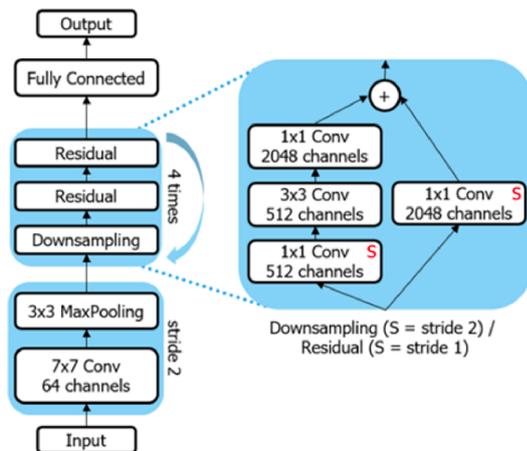


Figure 9. ResNet-50 Model

2.7.2. Non-pretrained models

The non-pretrained convolutional neural network (CNN) architecture utilized in this study consists of three 2D convolutional layers with 8, 16, and 32 filters, respectively. Each filter has a size of 3×3 , ensuring that each convolution operation processes a $3 \text{ pixel} \times 3 \text{ pixel}$ region of the input. The network also includes three pooling layers, three dense layers with 64, 32,

and 3 neurons, respectively, and two dropout layers, each with a dropout rate of 15%.

Convolutional neural networks (CNNs) are a type of artificial intelligence algorithm based on multilayer neural networks. These networks are designed to learn and extract relevant features from images, as illustrated in Figure 10. CNNs are capable of performing various tasks, including object classification, detection, and segmentation [10]. They are a fundamental component of the field of deep learning [11].

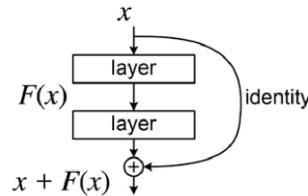


Figure 10. CNN

The Principal Component Analysis (PCA) model was also utilized in this study. PCA is a highly effective statistical technique widely applied in fields such as facial recognition and image compression. It is commonly used to identify patterns in high-dimensional data [12].

The ReLU (Rectified Linear Unit) activation function was employed in the convolutional neural networks (CNNs) used in this study. Its primary role is to enhance the nonlinear activation properties of the network without altering the receptive fields of the convolutional layers [13].

Convolutions

A convolution in an image is a pixel-by-pixel transformation achieved by applying a specific operation defined by a set of weights, commonly referred to as a filter. The convolutional layer in a neural network consists of a collection of learnable filters. Each filter is spatially small in terms of width and height but extends across the entire depth of the input volume [14].

Submapping

The pooling layer, also referred to as the subsampling layer, serves to progressively reduce the spatial dimensions of the representation, as illustrated in Figure 11. This reduction minimizes the number of parameters and computational complexity within the network [14].

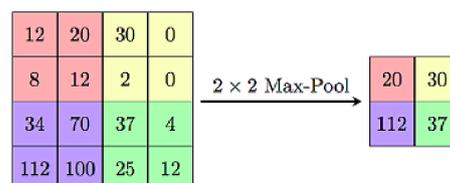


Figure 11. Submapping

Pooling layer

The pooling layer is utilized to reduce the dimensions of the feature maps, with the primary objective of decreasing processing times while preserving the most critical information. This dimensionality reduction helps mitigate overfitting in the network and introduces a degree of translation invariance [15].

Retinographies

Retinography is a diagnostic procedure that captures a non-invasive, painless color image of the fundus of the eye [16].

How CNNs work

Convolutional Neural Networks (CNNs) operate through machine learning [17] and supervised learning [18], leveraging several key components that function in an integrated manner. The core of CNNs lies in their convolutional layers, which perform convolution operations to analyze input images using small filters (kernels). These filters extract relevant features, such as edges, textures, and patterns, through matrix multiplication. By sliding across the image, the filters generate convolutional feature maps [19].

Activation functions

Following the convolution operation, a nonlinear activation function, such as the Rectified Linear Unit (ReLU), is applied. This introduces nonlinearity into the model, enabling it to capture and extract more complex features.

Pooling layers

These layers are employed to reduce the dimensionality of feature maps by summarizing the information extracted by the convolutional layers. This operation is typically performed using techniques such as max pooling or average pooling, which effectively reduce the size of the features while retaining their most relevant information.

Fully connected layers

After passing through multiple convolutional and pooling layers, the extracted information is flattened and fed into one or more dense (fully connected) layers. These layers perform classification or regression operations to generate the final output.

Regularization

To prevent overfitting and enhance the generalization capabilities of the model, regularization techniques are employed. These include methods such as dropout,

which randomly deactivates neurons during training to reduce reliance on specific features, and batch normalization, which normalizes the activations of intermediate layers.

Loss function and optimization

During training, a loss function is employed to quantify the discrepancy between the model's predictions and the actual labels. Optimization algorithms, such as stochastic gradient descent (SGD) and its variants, are then used to minimize this loss. By iteratively adjusting the weights of the neural network, these algorithms enhance the model's performance and predictive accuracy.

3. Results and discussion

ResNet (Residual Networks) addresses degradation issues in deep neural networks by introducing residual blocks. The primary differences among ResNet models lie in their depth, the size of the residual blocks, learning capacity, and computational cost. The training process was conducted in two phases, incorporating both the pre-trained ResNet models and the non-pre-trained CNN. This was performed using a dataset with imbalanced class distributions.

Phase-1

As shown in Table 2, during the training of the ResNet-18 model, the loss on the training set was observed to be 86%, while the validation loss (val_loss) was significantly higher, reaching 194%. This was accompanied by an accuracy of 60% and a validation accuracy (val_accuracy) of 70%. These results indicate potential calibration issues, early stopping, or improper training configurations caused by factors such as underfitting, excessive regularization, non-representative data, or sampling problems. For the ResNet-50 model, the training loss was 132%, and the validation loss was 126%, with training accuracy at 48% and validation accuracy at 54%. These metrics suggest challenges related to the model's learning and generalization capacity, possibly due to its increased complexity and computational requirements. In contrast, the non-pre-trained CNN demonstrated superior performance, achieving a training loss of 19% and a validation loss of 22%, with training and validation accuracies of 92% and 91%, respectively. The alignment of loss and accuracy metrics between the training and validation sets indicates that this model is generalizing well and effectively learning from the data.

As shown in Figure 12, a significant number of samples were classified into class 0 (No DR) with a count of 330. Class 1 (Mild) contained 19 samples, class 2 (Moderate) included 87 samples, class 3 (Severe) had

20 samples, and class 4 (Proliferative) comprised 35 samples.

Table 2. Training and validation phase-1

	ResNet-18	ResNet-50	CNN
Loss in training (loss)	0.86	1.32	0.19
Training precision (accuracy)	0.6	0.48	0.92
Loss in validation (val_loss)	1.94	1.26	0.22
Validation precision (val_accuracy)	0.7	0.54	0.91

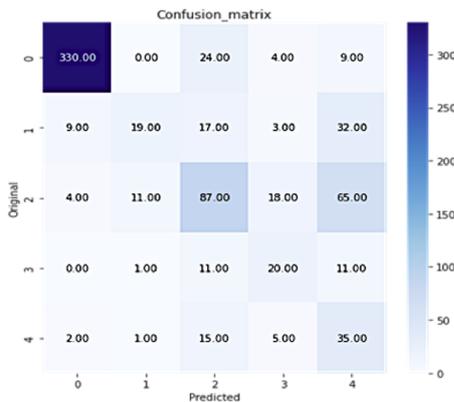


Figure 12. Confusion matrix ResNet18

Phase-2

In Phase-2, a series of adjustments were made to the hyperparameter configurations of the ResNet-18, ResNet-50, and non-pretrained CNN models to develop a robust and consistent model. As indicated in Table 3, the ResNet-based models showed no notable improvements compared to the results obtained in Phase-1. In contrast, the non-pretrained CNN model demonstrated significant enhancement in performance and precision, achieving an accuracy of 94%, a validation accuracy (val_accuracy) of 93%, a loss of 18%, and a validation loss (val_loss) of 19%. These metrics indicate effective generalization of the acquired knowledge, with consistent and reliable results. The non-pretrained CNN model clearly outperformed the ResNet-based models and proved to be a superior and more suitable choice for predicting Hepatic Retinopathy.

Table 3. Training and validation phase-2

	ResNet-18	ResNet-50	CNN
Loss in training (loss)	0.83	0.065	0.18
Training precision (accuracy)	0.68	0.48	0.94
Loss in validation (val_loss)	1.88	0.2	0.19
Validation precision (val_accuracy)	0.26	0.83	0.93

As depicted in Figure 13, a large number of samples were classified into class 0 (No DR), with a count of 351. Class 1 (Mild) included 8 samples, class 2 (Moderate) comprised 162 samples, class 3 (Severe) had 25 samples, and class 4 (Proliferative) accounted for 23 samples.

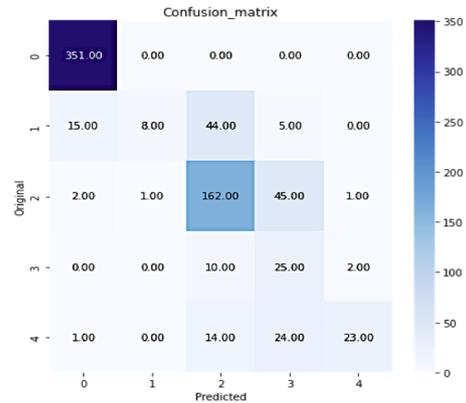


Figure 13. Confusion matrix ResNet50

The results obtained from the ResNet-based models raise several issues for discussion. The high loss observed in both training and validation phases may be attributed to the class imbalance within the dataset. Additionally, the combination of low accuracy values and high loss suggests that the models are not learning effectively from the data. This could be due to a lack of convergence or suboptimal hyperparameter configurations, leading to underfitting. Although ResNet-50 is inherently more powerful than ResNet-18 due to its greater depth and capacity, it may not be adequately suited or sufficiently tailored to the specific problem at hand.

The loss and accuracy indicators observed in Phase-1 and Phase-2 underscore the effectiveness of the non-pretrained CNN model. The high accuracy in both the training and validation sets suggests that the model successfully captures the patterns within the data and generalizes the acquired knowledge effectively. The minimal discrepancy between the training accuracy and validation accuracy (val_accuracy) is within acceptable limits and may be attributed to noise in the data or slight variations between the training and validation datasets.

The proposed approach for diabetic retinopathy detection offers significant advantages through its rigorous benchmarking of multiple neural network architectures. This process provides a comprehensive understanding of how various artificial intelligence models address a complex medical problem. The methodology is particularly notable for its ability to highlight the strengths and limitations of each architecture, demonstrating that increased model complexity does not necessarily translate into superior performance. The non-pretrained CNN emerged as a highly innovative

solution, achieving consistent accuracy exceeding 90%, robust generalization capabilities, and efficient handling of class imbalance key factors in diagnosing diseases characterized by rare but potentially severe presentations.

Despite its strengths, the proposed approach has notable limitations that warrant consideration. The reliance on a specific neural network architecture may restrict the transferability of the solution to other medical contexts, as the design is highly tailored to the dataset used in this study. Furthermore, the research highlighted the challenges faced by pre-trained models, such as ResNet-18 and ResNet-50, in adapting to medical datasets with complex and intricate features. This underscores the need for additional strategies, including advanced resampling techniques, weighted loss functions, and the augmentation of domain-specific data. These complexities introduce a more labor intensive development process, necessitating specialized expertise in both machine learning and the specific medical domain.

4. Conclusions

This analysis provided critical insights into the performance of various artificial intelligence models for the detection of diabetic retinopathy, highlighting significant variability among the evaluated neural network architectures [20].

ResNet-18 demonstrated critical limitations, with accuracy declining dramatically from an initial 70% to 26% in the final phase, underscoring its inadequacy for handling the complexity of medical image classification. In contrast, ResNet-50 exhibited a more robust learning capacity, achieving substantial improvement and reaching 83% accuracy in the final phase, emphasizing the importance of tuning and adaptation.

The non-pretrained CNN emerged as the most effective solution, consistently maintaining high levels of accuracy, nearing 91%, across both training phases and significantly outperforming the pre-trained models. This architecture achieved a training accuracy of 92% and a validation accuracy (val_accuracy) of 91% from the outset. Its stability across metrics and low validation loss (val_loss: 0.19 in Phase 2) demonstrated its capability to capture the necessary patterns for accurate image classification [21]. These results highlight that a carefully designed, simpler architecture can outperform more complex models in terms of efficiency and accuracy for specific problems.

Class imbalance was identified as a critical factor, particularly affecting the performance of the pre-trained ResNet models. The non-pretrained CNN handled this challenge remarkably well, suggesting that thoughtful architectural design can overcome the structural limitations of more complex models. While the

non-pretrained CNN successfully managed class imbalance, ResNet-18 and ResNet-50 struggled, particularly during the early training phases. This emphasizes the importance of implementing additional strategies, such as weighted loss functions, data augmentation, or advanced resampling techniques, to mitigate the impact of imbalance and enhance the performance of more complex models. Ensuring high-quality retinography images [?] is also crucial to avoid inconsistencies during the training phase.

Future research should focus on advanced strategies to manage class imbalance in medical datasets, addressing one of the most significant challenges identified in this study. These efforts should aim to create methodologies that ensure a more balanced representation of different image categories, particularly for minority classes that are critical to diagnosing diabetic retinopathy.

Proposed strategies include developing advanced resampling techniques, such as SMOTE, designing custom loss functions that dynamically weight classes, and creating data augmentation methods specifically tailored to medical images. These approaches aim not only to enhance model accuracy but also to improve their ability to detect rare yet clinically significant cases, representing a substantial advancement in the application of artificial intelligence to medical diagnosis.

The relevance of this work lies in its potential to transform AI systems capabilities for handling complex and imbalanced datasets, particularly in medical contexts where early and accurate detection is crucial for effective treatment. This direction offers promising avenues for improving diagnostic precision and addressing critical challenges in medical imaging.

References

- [1] OMS, *TADDS* Instrumento para la evaluación de los sistemas de atención a la diabetes y a la retinopatía diabética*. Organización Mundial de la Salud, 2015. [Online]. Available: <https://upsalesiana.ec/ing33ar8r1>
- [2] L. García Ferrer, M. Ramos López, Y. Molina Santana, M. Chang Hernández, E. Perera Miniet, and K. Galindo Reydmond, “Estrategias en el tratamiento de la retinopatía diabética,” *Revista Cubana de Oftalmología*, vol. 31, pp. 90–99, 03 2018. [Online]. Available: <https://upsalesiana.ec/ing33ar8r2>
- [3] F. Tablado. (2020) Inteligencia artificial en el trabajo ¿cómo afecta la ia al ámbito laboral de las empresas? Grupo Artico34. [Online]. Available: <https://upsalesiana.ec/ing33ar8r3>

- [4] E. W. Steyerberg, A. J. Vickers, N. R. Cook, T. Gerds, M. Gonen, N. Obuchowski, M. J. Pencina, and M. W. Kattan, "Assessing the performance of prediction models: a framework for some traditional and novel measures," *PubMed Central*, vol. 21, no. 1, pp. 128–138, 2010. [Online]. Available: <https://doi.org/10.1097/EDE.0b013e3181c30fb2>
- [5] L. Rouhiainen, *Inteligencia Artificial: 101 cosas que debes saber hoy sobre nuestro futuro*. Editorial Alienta, 2018. [Online]. Available: <https://upsalesiana.ec/ing33ar8r5>
- [6] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015. [Online]. Available: <https://doi.org/10.1038/nature14539>
- [7] J. M. Sanchez Muñoz, "Análisis de calidad cartográfica mediante el estudio de la matriz de confusión," *Pensamiento Matemático*, vol. 6, no. 2, pp. 9–26, 2016. [Online]. Available: <https://upsalesiana.ec/ing33ar8r13>
- [8] X. Ou, P. Yan, Y. Zhang, B. Tu, G. Zhang, J. Wu, and W. Li, "Moving object detection method via resnet-18 with encoder–decoder structure in complex scenes," *IEEE Access*, vol. 7, pp. 108 152–108 160, 2019. [Online]. Available: <https://doi.org/10.1109/ACCESS.2019.2931922>
- [9] G. Celano, "A resnet-50-based convolutional neural network model for language id identification from speech recordings," *ACL Anthology*, 2021. [Online]. Available: <https://upsalesiana.ec/ing33ar8r8>
- [10] E. Todt and B. A. Krinski, *Convolutional Neural Network–CNN*. Universidade Federal do Paraná, 2019. [Online]. Available: <https://upsalesiana.ec/ing33ar8r9>
- [11] J. Torres, *DEEP LEARNING Introducción práctica con Keras*. LuLu, 2018. [Online]. Available: <https://upsalesiana.ec/ing33ar8r10>
- [12] C. Zhu, C. U. Idemudia, and W. Feng, "Improved logistic regression model for diabetes prediction by integrating pca and k-means techniques," *Informatics in Medicine Unlocked*, vol. 17, p. 100179, 2019. [Online]. Available: <https://doi.org/10.1016/j.imu.2019.100179>
- [13] C. Bonilla Carrion, *Redes Convolucionales*. Trabajo Fin de Grado Inédito. Universidad de Sevilla, 2020. [Online]. Available: <https://upsalesiana.ec/ing33ar8r12>
- [14] D. Anguita, L. Ghelardoni, A. Ghio, L. Oneto, and S. Ridella, "The 'k' in k-fold cross validationthe 'k' in k-fold cross validation," in *European Symposium on Artificial Neural Networks, Computational Intelligenceand Machine Learning*, 2012. [Online]. Available: <https://upsalesiana.ec/ing33ar8r15>
- [15] L. Herguedas Fenoy, *Guía práctica clínica para la realización de una retinografía*. Universidad de Valladolid, 2018. [Online]. Available: <https://upsalesiana.ec/ing33ar8r17>
- [16] J. Chua, C. X. Y. Lim, T. Y. Wong, and C. Sabanayagam, "Diabetic retinopathy in the asia-pacific," *Asia-Pacific Journal of Ophthalmology*, vol. 7, no. 1, pp. 3–16, 2018. [Online]. Available: <https://doi.org/10.22608/APO.2017511>
- [17] G. Stiglic, P. Kocbek, N. Fijacko, M. Zitnik, K. Verbert, and L. Cilar, "Interpretability of machine learning-based prediction models in healthcare," *WIREs Data Mining and Knowledge Discovery*, vol. 10, no. 5, p. e1379, 2020. [Online]. Available: <https://doi.org/10.1002/widm.1379>
- [18] Z.-H. Zhou, "A brief introduction to weakly supervised learning," *National Science Review*, vol. 5, no. 1, pp. 44–53, Jan 2018. [Online]. Available: <https://doi.org/10.1093/nsr/nwx106>
- [19] J. Wu, *Introduction to Convolutional Neural Networks*. LAMDA Group, 2017. [Online]. Available: <https://upsalesiana.ec/ing33ar8r22>
- [20] R. Gonzalez Gouveia, *Diferencias entre Inteligencia Artificial vs Machine Learning vs Deep Learning*. YouTube, 2021. [Online]. Available: <https://upsalesiana.ec/ing33ar8r23>
- [21] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," *arXiv*, 03 2015. [Online]. Available: <https://upsalesiana.ec/ing33ar8r24>



DESIGN OF A EDUCATIONAL ELECTRICAL ENERGY CONSUMPTION METER FOR RESIDENTIAL USE

DISEÑO DE UN MEDIDOR DIDÁCTICO DE CONSUMO DE ENERGÍA ELÉCTRICA PARA USO RESIDENCIAL

Alan Cuenca Sánchez^{1,*} , Pablo Llumiquinga Eras¹ 

Received: 13-09-2024, Received after review: 25-11-2024, Accepted: 02-12-2024, Published: 01-01-2025

Abstract

This study presents the design and implementation of an interactive, user-friendly electronic meter for energy consumption measurement. The proposed system serves as an educational tool for teaching electrical installations, offering a practical and hands-on learning experience. The primary objective is to develop an interactive meter tailored for residential use, capable of providing real-time feedback on energy consumption. Its deployment in educational institutions enhances the comprehension of technical concepts, while also proving beneficial in community outreach workshops focused on residential electrical systems. The system consists of a battery-powered setup featuring an ESP32 module for voltage and current data acquisition, an SPI-connected LCD screen for local data visualization, and a WiFi module for real-time data transmission to a cloud-based database. Designed to be reproducible, cost-effective, and open source, the system represents an accessible and versatile solution for energy monitoring applications. Validation tests were conducted over five months in both laboratory and residential environments. The results demonstrated high measurement accuracy, with error margins below 5% for voltage, current, energy consumption, and estimated costs. These findings confirm that the developed interactive energy meter is a reliable and effective tool for monitoring residential energy usage while fostering educational and community-based learning experiences.

Keywords: database, consumption, energy, electric energy meter, residential

Resumen

Este trabajo presenta el diseño e implementación de un sistema de medición de consumo de energía eléctrica mediante un medidor electrónico interactivo y amigable para el usuario. Este sistema constituye una herramienta didáctica para la enseñanza de instalaciones eléctricas, ofreciendo una experiencia práctica y educativa. El objetivo principal es desarrollar un medidor interactivo para uso residencial que proporcione retroalimentación en tiempo real sobre el consumo energético. Su implementación en instituciones educativas no solo facilita la comprensión de conceptos técnicos, sino que también resulta valiosa en talleres prácticos de vinculación social enfocados en instalaciones eléctricas residenciales. El sistema consta de una batería, un módulo ESP32 para adquirir datos de voltaje y corriente, una pantalla LCD conectada por SPI para visualización local y un módulo wifi que transmite los datos a una base de datos en la nube. Su diseño es reproducible, económico y de código abierto, lo que lo convierte en una solución accesible y versátil. Pruebas realizadas durante cinco meses en laboratorios y domicilios validaron su precisión, con márgenes de error menores al 5% en voltaje, corriente, energía y costos estimados. Los resultados confirman que este medidor interactivo es una herramienta eficaz, viable e interactiva para el monitoreo energético en entornos residenciales.

Palabras clave: base de datos, consumo, energía, medidor de energía eléctrica, residencial

^{1,*}Escuela de Formación de Tecnólogos, Escuela Politécnica Nacional, Ecuador.
Corresponding author ✉: alan.cuenca@epn.edu.ec.

Suggested citation: Cuenca Sánchez, A. and Llumiquinga Eras, P. "Design of a educational electrical energy consumption meter for residential use," *Ingenius, Revista de Ciencia y Tecnología*, N.º 33, pp. 102-114, 2025, DOI: <https://doi.org/10.17163/ings.n33.2025.09>.

1. Introduction

Promoting the efficient use of electricity is a fundamental component of a country's development plans. Achieving this goal requires effective demand management, which relies on accurately measuring and monitoring electricity consumption by end users [1].

In Ecuador, energy savings in the residential sector remain low. Although some energy efficiency plans have been implemented, there are no permanent or long-term initiatives in place [2]. Nonetheless, a significant portion of the population recognizes the importance of energy efficiency [3], particularly due to its potential for reducing electricity costs and achieving economic savings [4].

Energy-saving practices are often hindered by the misinterpretation of utility bills [5]. Moreover, these bills typically lack real-time feedback mechanisms that enable consumers to evaluate the effectiveness of their energy-saving efforts [6]. Consequently, even conscientious consumers are unable to receive immediate feedback to fully recognize and appreciate the benefits of adopting improved electricity consumption habits [7].

Technological advances in the design and construction of energy measurement systems have significantly progressed. These advancements have led to the development of electronic meters that record energy consumption exclusively in kilowatt-hours (kWh). However, such meters often present challenges for consumers to interpret and fail to effectively raise awareness of electricity usage [8].

Currently, smart meters provide more detailed data than conventional meters [9]. This enhanced information enables consumers to develop better energy consumption habits and facilitates the implementation of effective energy-saving techniques [10].

Measurement based on continuous monitoring and data recording provides users with comprehensive insights into energy consumption, including details on usage quantities, patterns, and associated electricity costs. This approach facilitates optimal resource management, identifies system losses, and enables targeted actions to improve household energy efficiency [11].

In response to these challenges, an interactive electric energy meter is being developed for the residential sector. The meter is designed to provide real-time information to users through a graphical interface displayed on an LCD screen. The presented data include energy consumption, voltage levels, electric current, power, and estimated monthly billing costs, with a maximum measurement error of 5%. Furthermore, the system incorporates a cloud-based database that stores hourly historical data on electric energy consumption.

This meter is an essential educational tool for promoting knowledge of residential electrical installations. It facilitates the development of more advanced prototypes with enhanced functionalities, ultimately con-

tributing to the efficient use of electrical energy.

1.1. State of the art

Estrada [12] designed and developed a prototype for measuring energy consumption, which was validated in a laboratory setting by collecting data at 3-minute intervals over a 4-hour period. The study highlighted the significance of providing households with dynamic and interactive energy consumption data. Additionally, it emphasized the limitations of current residential meters, which require manual readings by an operator. This process often introduces accuracy errors, leads to suboptimal service quality, and results in dissatisfaction among residential users.

Samaniego and Velezaca [13] implemented an electronic energy meter for residential use, noting that a significant percentage of users paying for electricity services each month are unaware of the rates charged by the distribution company. The proposed meter addresses this issue by sending text messages to users, ensuring they remain consistently informed about their energy consumption and associated costs.

Vashist and Tripathi [14] developed and implemented a smart energy meter integrated with real-time pricing in an Android-based web application. However, the system lacks daily or monthly measurement capabilities, which limits the assessment of the device's accuracy. Similarly, Patel et al. [15] designed and deployed a smart meter with load-forecasting capabilities for residential customers. The system utilizes GSM communication and transmits AT commands via a microcontroller. However, the operators noted that this approach may incur higher costs due to the transmission of messages.

In [16], the design and implementation of an electronic meter are described, enabling real-time monitoring of electricity consumption. The system allows users to visualize both current and historical consumption through a graphical interface on a computer. However, this study focuses on more generalized residential solutions, in contrast to the device presented in this article, which employs specific technologies such as ESP32 and Firebase for cloud-based data logging.

The study conducted by Muñoz et al. [17] focuses on the design and implementation of an integrated system for both local and remote monitoring of electricity consumption across various areas of a household. This system also measures the current consumed at one-minute intervals by different loads connected to the domestic electrical grid. In contrast, the prototype presented in this work employs open-source technology and is portable, cost-effective, and easier to replicate in educational laboratories. Moreover, studies presented in [18] and [19] examine the implementation of advanced metering systems, commonly referred to as smart meters, and address the challenges associated

with promoting responsible electricity consumption. These findings align with the objectives of this study, underscoring the importance of accurate measurement as a fundamental tool for optimizing energy consumption and fostering energy-saving practices through historical data analysis.

Morales and Peña [20] present the design and implementation of a Home Energy Management (HEM) system within the energy sector in Colombia. The system monitors the energy consumption of typical household devices, allowing users to analyze individual device usage and develop strategies to reduce energy consumption at home. While both the HEM system and the prototype proposed in this work promote the adoption of monitoring technologies to encourage energy savings, the interactive meter presented here offers notable advantages: it is cost-effective, user-friendly, and easy to build. Additionally, its integration of internet connectivity enhances its applicability in social and educational contexts, significantly contributing to the academic and practical development of both students and end users.

The studies reviewed above emphasize the importance of household energy consumption measurement systems. However, none propose the design and implementation of a cost-effective system that incorporates energy backup, is easy to replicate, and is simple to operate, while offering real-time data access both locally (via a touchscreen interface) and digitally (through a database). The system developed in this work is intended for academic purposes and has been tested and validated in a real household over a five-month period. It aims to serve as a foundation for community outreach projects and practical workshops, enabling participants to understand system operation, gain insight into how electricity services are billed in Ecuador, and adopt measures to promote energy efficiency and savings. Furthermore, this study provides the complete technical documentation required to build the meter, underscoring its contribution to the field of residential electrical installations.

The system developed in this research can be utilized in classrooms as an educational tool to teach students about residential electrical installations, serving as a practical example of a didactic monitoring system.

Table 1 presents several studies on electricity consumption meters, examining the methods, error correction techniques, costs, platforms, and technologies employed. These studies highlight the importance of implementing such meters in households, even with basic functionalities, such as real-time reporting of consumption and cost data, to encourage energy-saving behaviors and promote the rational use of energy by consumers.

Table 1. Related studies on electricity consumption meters in the residential sector

Article	Analyzed characteristic
[21]	Connectivity
[22]	Monitoring method
[23]	Platforms
[24]	Technology
[25]	Evolution

Arévalo et al. [26] highlight that limited research has focused on analyzing the energy consumption habits of residential users using digital meters that provide real-time access to historical consumption data. This capability enables users to take informed actions to effectively reduce their electricity consumption. Similarly, Alahmad et al. [27] emphasize that altering consumer behavior plays a critical role in reducing residential energy usage. Their study examines the impact of digital meters on residential energy consumption rates in a metropolitan area, raising awareness among residents about their electricity consumption patterns and the environmental benefits of energy savings.

Building on these findings, this research aims to design, develop, and validate a cost-effective and easily replicable didactic system for measuring and recording voltage, current, energy consumption, and electricity billing costs in households. The system is designed to facilitate knowledge dissemination in the field of residential electrical installations. To achieve this, low-cost sensors are employed to monitor the household electrical system, enabling the acquisition of accurate energy consumption data. The obtained values are compared with those recorded by conventional meters and measurement instruments, such as multimeters, to verify the margin of error and ensure the system's accuracy and reliability. The system developed in this study can be further enhanced to improve data accessibility and reduce costs in similar projects [12–17]. Additionally, it holds significant educational potential as a practical tool in university laboratories and educational centers for teaching residential electrical installations. By enabling students to understand the operation of interactive, user-friendly energy consumption meters, the system fosters the development of strategies to promote energy efficiency in households.

2. Materials and methods

This section outlines the stages of the project development, detailing the materials and methods employed. The proposed system is divided into two main components:

1. Electrical-Electronic system
2. Data measurement and logging system

Figure 1 illustrates the primary components of the energy meter developed in this study.

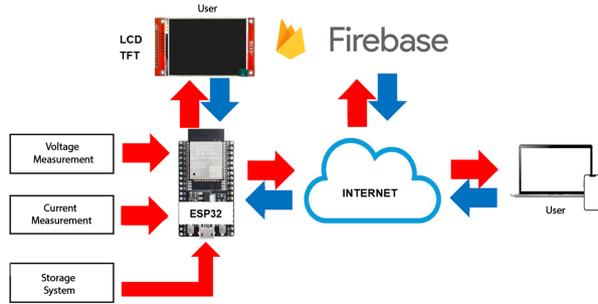


Figure 1. Primary components of the residential didactic energy meter

2.1. Electrical-Electronic system

The meter’s power supply was designed with an integrated protection system utilizing a fuse. The design features a transformer that steps down the voltage from 120 VAC to 12 VAC, paired with a diode bridge for rectification and battery charging. A relay automatically selects the system’s power source, seamlessly switching to the battery as the primary power supply during a power outage. Capacitors ensure a temporary voltage supply during this transition, maintaining uninterrupted system operation. Figure 2 illustrates the electrical schematic of the power supply and charging system.

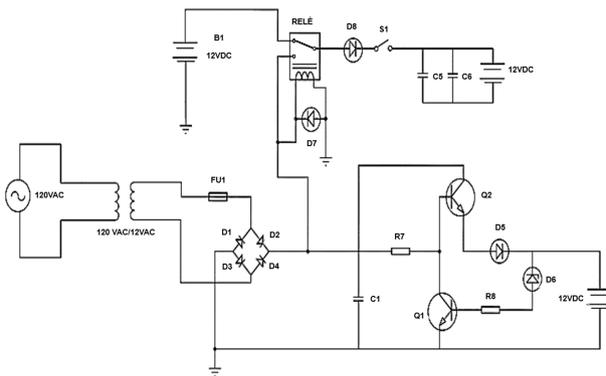


Figure 2. Power supply circuit

The design of the conditioning circuits relies on the EmonLib library [28], which computes active power, apparent power, power factor, RMS voltage, and RMS current using discrete-time equations for these parameters. Therefore, preserving the alternating signal after conditioning is essential to ensure accurate measurements.

The alternating voltage and current signals are sampled by the microcontroller, with both signals obtained through transformers. The conditioning circuits

for these signals share a similar design. For the voltage signal, a voltage divider is used to scale it within the limits permitted by the analog-to-digital converter (ADC) inputs of the microcontroller. A second voltage divider introduces a 1.65 VDC offset to the scaled alternating signal to prevent negative values during voltage measurement. This offset corresponds to half of the maximum voltage allowed by the ADC inputs. Figure 3 illustrates the schematic of the voltage signal conditioning circuit.

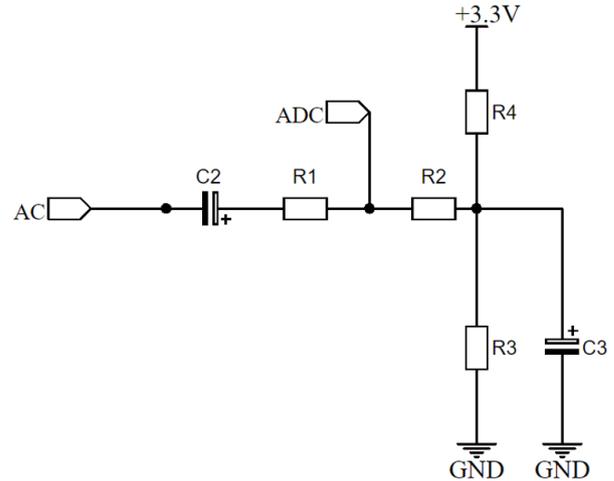


Figure 3. Voltage conditioning circuit

To determine the resistor values for the first voltage divider, Equation (1) is applied. The maximum output voltage of the divider (V_0) must remain within the allowable maximum limits of the ADC inputs while accounting for variations in the grid voltage (V_s).

$$V_0 = V_s \left(\frac{R_2}{R_1 + R_2} \right) \quad (1)$$

Where, V_s is the voltage on the secondary side of the transformer; R_1 is the fixed resistor of the voltage divider, V_0 is the output voltage of the divider, and R_2 is the resistor to be determined.

This circuit ensures that when the grid voltage reaches its maximum of 129.60 VAC, the microcontroller receives an alternating voltage ranging between 2.40 V (maximum) and 0.80 V (minimum). This design prevents negative voltages while preserving the integrity of the alternating signal.

For the current conditioning circuit shown in Figure 4, a load resistor is required on the secondary side of the current transformer to generate a voltage signal suitable for sampling. The value of this resistance is determined using Equation (2).

$$R_L = \frac{(N_{CT})(V_{CT})}{1} \quad (2)$$

Where, V_{CT} is the required voltage on the secondary side of the current transformer; I is the maximum current to be measured; N_{CT} is the transformation ratio; and R_L is the load resistance.

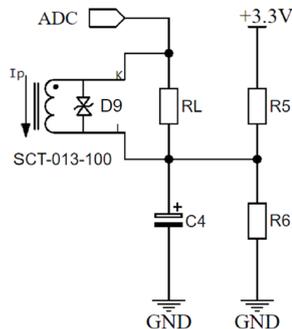


Figure 4. Current conditioning circuit

This voltage signal obtained across the load resistor is offset by 1.65 VDC using a voltage divider circuit, where resistors R_5 and R_6 each have a value of $30\text{ k}\Omega$. This configuration ensures that the voltage provided by the divider is half of the microcontroller's supply voltage.

The capacitor C_4 provides a low-impedance path, allowing the AC signal current to flow to ground (GND) without passing through R_6 . For a capacitance of $10\ \mu\text{F}$ and a frequency of 60 Hz, the impedance is calculated as $265.26\ \Omega$, which is significantly lower than the resistance of R_6 .

With the selected resistors and capacitors, the microcontroller is ensured to receive an alternating voltage ranging from a maximum of 2.65 V to a minimum of 0.65 V when measuring the peak current of 66.72 A obtained from the load study.

A 32-bit microcontroller, specifically the ESP32 development board, is employed due to its analog input and digital output capabilities, which enable seamless integration with various components. Additionally, the ESP32 features an integrated circuit for internet connectivity. This microcontroller processes the voltage and current signals, performs the necessary calculations to determine the corresponding electrical parameters, and controls the LCD screen to display the relevant data. Furthermore, it transmits the information to a cloud database via WiFi. As a 32-bit system, the ESP32 provides enhanced precision for floating-point calculations and includes a 12-bit analog-to-digital converter, offering improved resolution during the sampling of voltage signals.

With the circuit designs completed, the printed circuit board (PCB) design, shown in Figure 5, was created using Easy EDA design software. The design incorporates several considerations, including connections for the TFT screen, a ground plane to minimize electromagnetic interference (EMI), the avoidance of 90° bends in traces, and adherence to standard sizes

for electronic components. It is important to note that, due to the type of connection required for the LCD screen and its touch functionality, the same SPI port of the microcontroller was utilized. Consequently, cable bridges were implemented on the top layer, indicated in red in the PCB design.

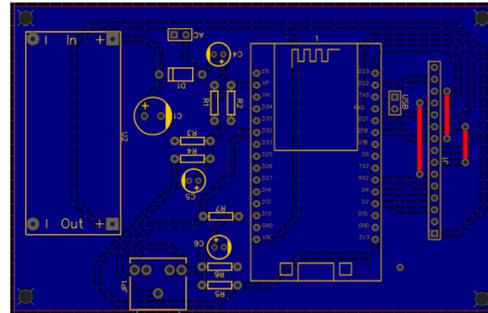


Figure 5. PCB Design

The design of the meter case was developed considering the dimensions and arrangement of all required components, including the transformer, battery, electronic board, power button, relay, screen, and holes for screws and wires. A suitable design was selected to meet the functional requirements of the meter and was created using 3D modeling in Fusion 360 software. Figure 6 illustrates the design of the meter case.

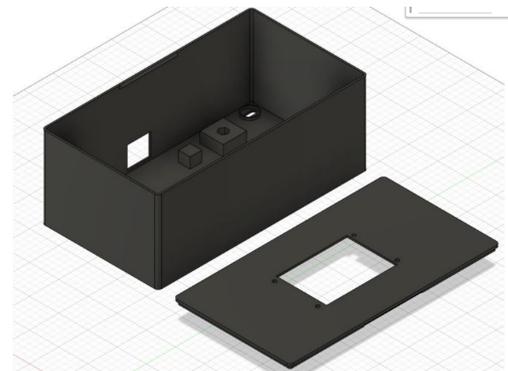


Figure 6. Design of the meter case

2.2. Data measurement and recording system

The measurement programming code was developed using the Arduino IDE, using multiple libraries essential for various processes, such as the calculation of electrical parameters, control of the TFT screen with touch functionality, internet connectivity, and connection to the Firebase database.

The microcontroller utilizes the current date and time, obtained from the internet, to record energy consumption values in the database corresponding to each hour of the day. Based on the measured energy data, the economic cost of consumption is calculated. This

data is reset every 30 days and archived in the historical records. The monetary calculations are performed according to the tariff schedule by consumption range provided by Empresa Eléctrica Quito [29].

The flowcharts detailing the measurement system, signal conditioning, and data recording processes are presented in Figures 7 and 8.

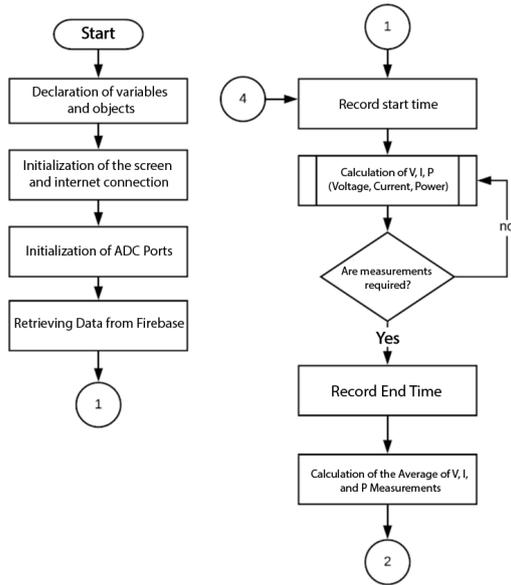


Figure 7. Flowchart for measurement and data recording system - Part 1

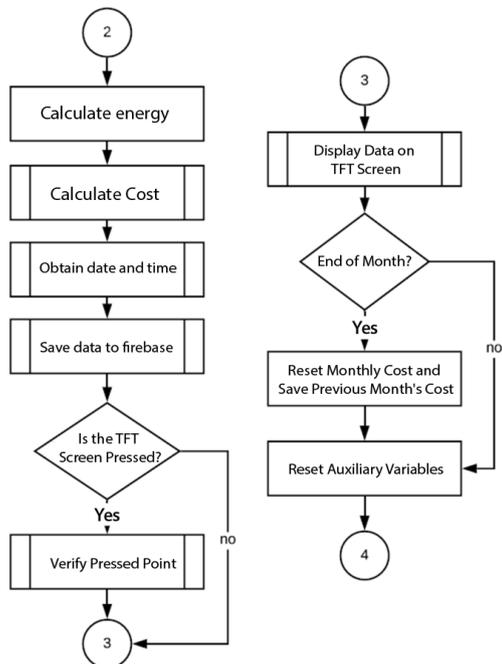


Figure 8. Flowchart for measurement and data recording system - Part 2.

Wi-Fi connectivity enables the meter to store recorded energy consumption data in a database. Key

parameters relevant to the user, including power, voltage, current, consumed energy, and energy cost, are logged and stored. In addition to providing real-time feedback on electrical variables, the database remains accessible to users at any time from any internet-connected device. The primary advantage of using wireless communication, particularly Wi-Fi, is the ability to monitor measurements in real time. To access the database, users must have a Gmail account, which allows them to be added as viewers in the Firebase project. Once registered, a unique URL is provided, granting access to the real-time database via any web browser on a mobile device or computer, enabling users to monitor their energy consumption conveniently.

3. Results and discussion

Considering the design parameters and characteristics of the energy meter for residential use, a didactic, cost-effective, and easily reproducible system was developed. The system interactively displays real-time consumption data to the user, both on a local TFT screen and through an online data log accessible from any internet-connected device.

The printed circuit board (PCB) was fabricated using the CNC method on fiberglass material, while the meter case was constructed via 3D printing.

Once all components were prepared, the final assembly of the meter was carried out, integrating all elements within the case. A data bus was implemented for the TFT screen connection, ensuring quick and efficient assembly and disassembly. Figure 9 presents the completed meter, including its screen, which displays the corresponding real-time data.



Figure 9. Educational meter for residential energy consumption

To ensure the accurate measurement of electrical variables, the meter was installed near the load center of the residential home under study, as illustrated in Figure 10. This location centralizes energy distribution to the various circuits of the house, minimizing losses and ensuring representative measurements. The current transformer was connected in series with the main power phase to acquire a signal proportional to

the total system current. Meanwhile, the voltage transformer was connected between the phase and neutral terminals, providing a reduced and representative signal of the line voltage. Both signals were conditioned through circuits designed to adjust their amplitude and introduce an offset, ensuring they remain within the acceptable range of the microcontroller's analog-to-digital converter (ADC) inputs. This configuration ensures accurate data acquisition and enables subsequent processing for the calculation of power, energy, and other relevant parameters.

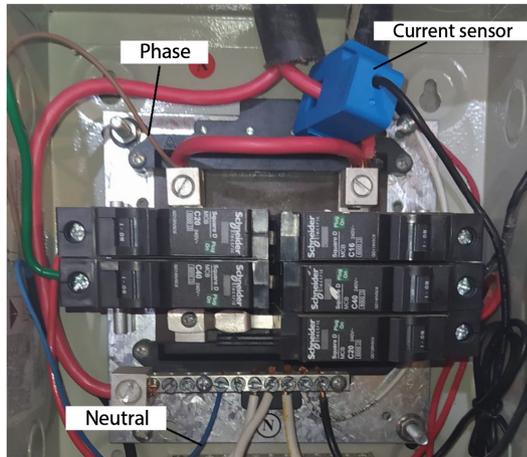


Figure 10. Electrical connections at the load center

For the calibration of the voltage and current sensors, measurements were performed using a multimeter as the reference for the residential grid voltage and a clamp ammeter for current measurements.

The calibration constants were iteratively adjusted within the measurement algorithm to minimize deviations between the meter readings and the reference instrument measurements.

Multiple voltage and current measurements were conducted at various times throughout the day. When discrepancies were identified between the readings from the multimeter and those from the meter, the voltage and current calibration constants were adjusted using Equation (3).

$$K_n = K_a \left(\frac{M_p}{M_m} \right) \quad (3)$$

Where, K_n is the new calibration constant, K_a is the previous calibration constant, M_p is the reading from the multimeter, and M_m is the reading from the meter.

The NTE INEN-IEC 62053-21 standard [30] specifies that, for residential systems, meters must conform to Class 1 or Class 2 requirements. These include a nominal operating frequency of 60 Hz, an LCD display with the number of integer digits determined by the requirements of the distribution companies, and an admissible error limit of 5% of the nominal values.

In the test household, daily measurements recorded by the conventional meter were documented. Additionally, an analysis of electricity bills from the past five months was conducted. Using this data, along with a load study, the average daily and monthly consumption was determined.

As presented in Table 2, the consumption values obtained from the three studies—load study, daily measurements, and electricity bills—are consistent, thereby validating the data and confirming the functionality of the constructed meter.

Table 2. Average electricity consumption of the test household

Type of consumption	Load Study (kWh)	Daily Measurement (kWh)	Electricity Bills (kWh)
Daily	6.84	6.62	6.32
Monthly	202.32	198.57	189.60

With the household's average energy consumption levels established, an analysis was conducted to determine the voltage and current levels the meter must measure. For voltage, the permitted variations in the residential sector, as specified by the Energy and Natural Resources Regulatory and Control Agency (ARCERNR), are $\pm 8\%$ [31]. Given the standard residential voltage of 120 VAC, the permissible range extends from a maximum of 129.60 VAC to a minimum of 110.40 VAC. The maximum current, on the other hand, depends on the household's total power demand at full load.

For the system configuration in complex residential environments, the following aspects are considered:

- **Robust design for electrical variations**

The device is engineered to withstand voltage fluctuations and measure peak currents up to 66.72 A. This design incorporates transformers paired with calibrated voltage dividers to ensure input signals remain within the microcontroller's operational limits. Conditioning circuits with offset adjustments are employed to prevent distortion of alternating signals. Additionally, the system includes a current transformer with a load resistance specifically calculated using equations based on the expected maximum power.

- **Modularity for different residential configurations**

The system is designed to be adaptable to standard single-phase residential systems and can be expanded to accommodate biphasic or triphasic configurations by adjusting the current and voltage sensors. Its compact design facilitates the integration of additional components, enabling the inclusion of more sensors or functions as needed.

- **Advanced sensor calibration**

In environments with frequent fluctuations, the system supports recalibration of voltage and current constants using adjustment formulas derived from reference measurements, ensuring long-term accuracy. Furthermore, continuous testing can be conducted under varying conditions to correct deviations by applying absolute and relative percentage error calculations, with acceptable limits set below 5%.

- **Real-time interaction and feedback**

The LCD screen provides real-time data on voltage, current, power, and costs, which is essential in residential environments where constant monitoring is required. Additionally, remote access to the database enables users to identify consumption patterns and detect anomalies from any internet-connected device.

- **Resilience to electrical interruptions**

The system is designed to remain operational during power outages, ensuring uninterrupted monitoring and data recording. The design incorporates a relay that automatically switches the battery and the main power supply.

- **Adaptability to regulations and standards**

The error limits, consumption rates, and recording frequencies can be adjusted to comply with regulatory requirements or to meet the specific needs of the residential area.

- **Scalability for complex residences**

The system offers the potential for implementing a network of interconnected meters to monitor entire residential buildings or areas with multiple housing units. Integration with IoT systems could enhance compatibility with home automation platforms, optimizing energy usage in conjunction with smart devices such as thermostats and solar panels. Furthermore, additional sensors could be incorporated to monitor variables such as electrical frequency and power quality, including harmonic distortion and voltage fluctuations.

3.1. Tests

Voltage and current tests were conducted by recording hourly measurements over a day (24 measurements). Figure 11 presents the measured voltage data. The comparison between the multimeter and the meter readings reveals an absolute and relative percentage error close to zero, with a mean absolute error (MAE) of 0.520. The mean absolute percentage error (MAPE) is also significant, with a value of 0.433%, well within the error tolerance of 5%. These results validate the

accuracy of the system's voltage measurements during the conducted tests.

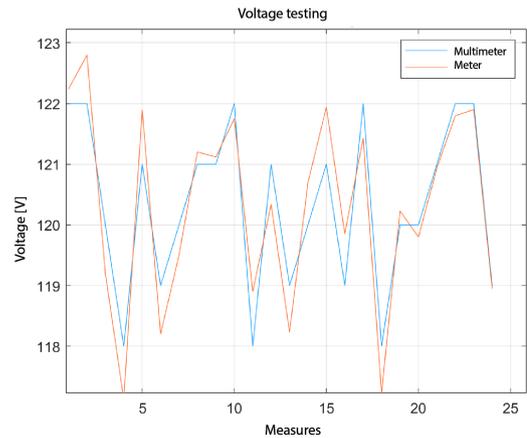


Figure 11. Voltage comparison between the multimeter and the proposed system

Figure 12 illustrates the measured current data. The comparison between the clamp meter and the system's meter readings reveals a mean absolute error (MAE) of 0.180. Additionally, the mean absolute percentage error (MAPE) is significant, with a value of 1.90%, which is well within the error tolerance of 5% as established by national regulations. These results confirm the accuracy of the system's current measurements during the conducted tests.

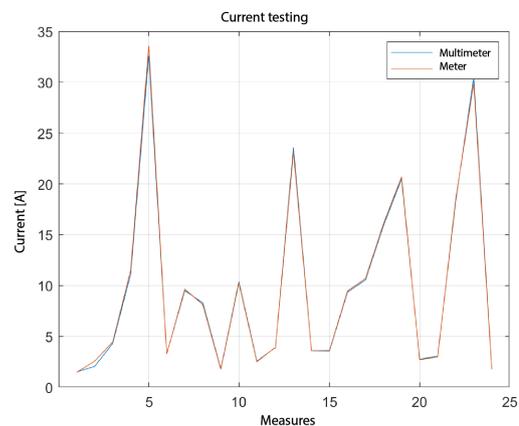


Figure 12. Current comparison between the clamp meter and the proposed system

After verifying the current and voltage measurements, the data storage process in the database was reviewed. It was confirmed that the information is stored daily in the desired format (DD-MM-YYYY) and that each record includes the corresponding hourly sub records (00:00–23:00). Additionally, it was verified that consumption values are updated hourly and reset at the start of each new day. The results are presented in Figures 13 and 14.

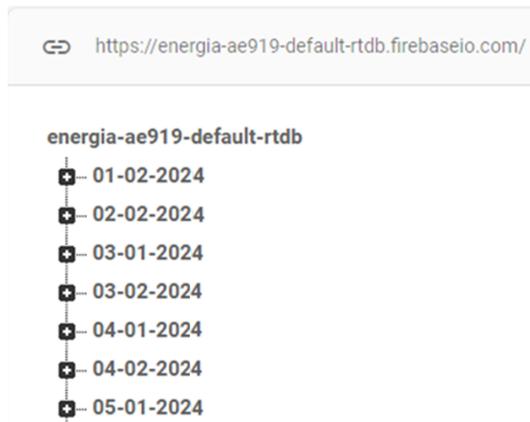


Figure 13. Daily records in Firebase



Figure 14. Hourly sub records in Firebase

Additionally, the accuracy of electrical energy measurement and cost calculation was validated. A comparison was conducted between the measurements obtained from the developed meter and those from a conventional meter, as well as between the energy costs calculated by the system and the economic values recorded on electricity bills over a five-month period.

Figure 15 presents the daily energy data measured by both the interactive meter and the conventional meter over a month, with measurements taken at the same time each day. Notably, unlike the developed interactive meter, the conventional meter does not report decimal values. Absolute and relative percentage errors were generally close to zero, with a mean absolute error (MAE) of 0.127 and a mean absolute percentage error (MAPE) of 1.90%. These results fall within the 5% error margin permitted by regulatory standards.

These results validate the performance of the developed meter, demonstrating its accuracy and precision in comparison with conventional instruments across various measurement parameters.

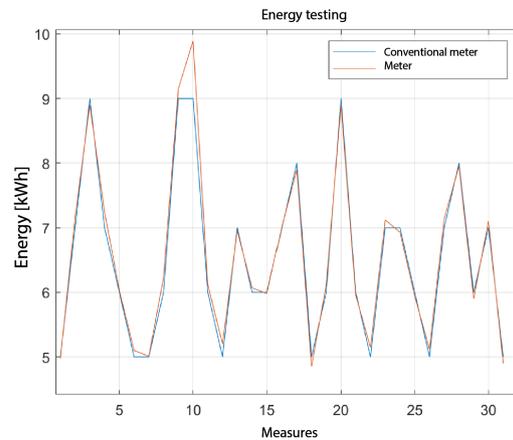


Figure 15. Daily energy comparison between the conventional meter and the proposed system

To verify the final operation of the interactive meter, energy consumption measurements recorded over 29 days in February 2024 were compared with the corresponding calculated economic values, as illustrated in Figure 16. These results were then contrasted with the data from the electricity bill for the same month, as shown in Figure 17.



Figure 16. Monthly record of the meter

Consumo Total	Unit of measure	Amount (\$)
201,00	kWh	17,11

Figure 17. Values provided on the electricity bill

The user interface includes a touch-enabled button that allows navigation between two information windows. The first window displays the monthly energy consumption in kilowatt-hours (kWh) and the corresponding cost in US dollars. The second window presents real-time voltage, current, power, and the monetary value associated with the previous month's consumption.

Finally, measurements were conducted over approximately five months (from March to July 2024) to verify

the reliability of the developed meter. The monthly energy data recorded by the meter, along with the corresponding values from the electricity bills, are presented in Figure 18 and Table 3.

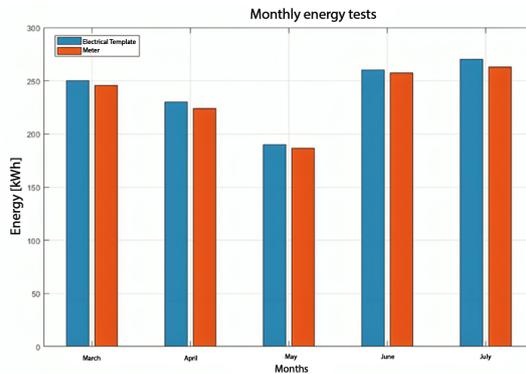


Figure 18. Monthly energy comparison between the electricity bill and the proposed system

Table 3. Comparison of monthly energy consumption

Months	Electricity bill (kWh)	Meter (kWh)	Absolute error (kWh)	Relative (error %)
March	250	247.8	2.2	0.88
April	230	228.8	1.2	0.52
May	190	188.4	1.6	0.84
June	260	257.6	2.4	0.92
July	270	267.3	2.7	1.00

Figure 19 and Table 4 present the monthly energy cost data calculated by the constructed meter alongside the values obtained from the electricity bills.

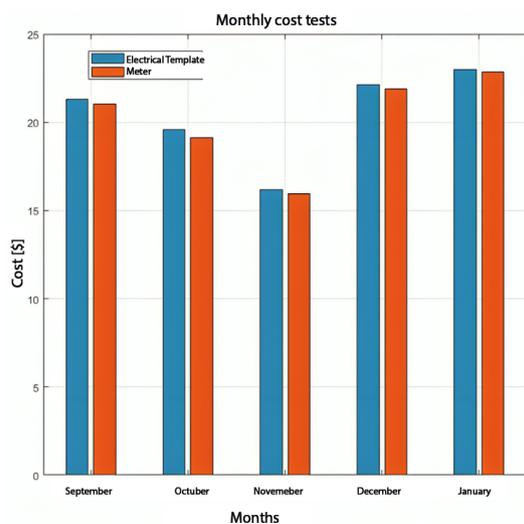


Figure 19. Monthly economic cost comparison between the electricity bill and the proposed system

Table 4. Comparison of energy consumption monetary cost

Months	Electricity bill (\$)	Constructed meter (\$)
March	21.28	21.02
April	19.57	19.12
May	16.17	15.96
June	22.13	21.90
July	22.98	22.85

The various tests conducted validate the proper functioning of the constructed meter for residential use, demonstrating its ability to provide accurate data and securely store it in the cloud. This ensures reliable information regarding electricity consumption and its associated economic cost.

The results obtained from tests conducted in households and at the Industrial Technology Laboratory of the ESFOT Technologist Training School at the National Polytechnic School demonstrate the effectiveness of the residential energy consumption meter. Its primary objective is to support the academic development and practical training of Electromechanics students. Figure 20 illustrates a technical workshop led by ESFOT students on residential electrical installations for young people from vulnerable communities, conducted as part of the outreach project with the 'Río Verde' Foundation. The meter (1) is used alongside the training teaching panels (2) to facilitate these workshops.



Figure 20. Students participating in technical workshops on residential electrical installations as part of outreach projects

4. Conclusions

An interactive electric energy meter for residential use was successfully developed and implemented, designed to deliver accurate measurements with an error margin of less than 5%, in compliance with national regulations. The device provides users with real-time information on their energy consumption, including estimated monthly costs, and maintains a historical record stored in the cloud. Equipped with a TFT LCD touchscreen, the meter displays essential data such as voltage levels, current, power, energy consumption

(kWh), and costs (USD), all of which are simultaneously recorded in an online database.

To extend historical records beyond those typically maintained by electricity distribution companies, the system incorporates an hourly consumption logging feature that requires an internet connection and synchronization with precise time and date data. The database stores daily energy consumption resets the counter monthly, and updates backup data to ensure accuracy. Firebase Real-Time Database was selected as the primary tool for its ability to store data in a simple, nested, and customizable manner, as well as being cost-free during the project's development. The collected data provides a valuable resource for analyzing residential consumption habits and detecting anomalies, thereby facilitating the adoption of energy-saving practices.

The developed meter offers enhanced functionalities compared to conventional meters by providing real-time data that enables users to monitor their energy consumption and design tailored strategies for cost reduction. While currently equipped with basic functions targeted at residential consumers, the system is flexible, adaptable, and cost-effective, with an approximate production cost of \$80 (USD). Its portability and ease of replication further increase its potential for widespread adoption. Future enhancements could include advanced features such as prepaid and postpaid modes, remote access for energy provider management, and fault detection capabilities.

This interactive residential energy meter significantly contributes to fostering sustainability by addressing issues related to efficient electricity use and waste reduction. Its ability to provide real-time data on energy consumption, estimated costs, and historical records empowers users to make informed decisions, optimize consumption habits, and reduce greenhouse gas emissions associated with energy waste. Additionally, its application in educational environments enhances technical knowledge of electrical installations and raises awareness about the importance of energy conservation in achieving environmental sustainability.

The designed and implemented device represents a significant advancement in the field of residential electrical installations and energy savings. It promotes learning and technical training through its integration into educational institutions, such as ESFOT, as part of the Applied Technology project line. This initiative addresses critical technical challenges in the country and finds practical applications in electricity laboratories and training programs for residential electrical installations. Furthermore, its incorporation into community outreach projects, such as technical workshops, fosters skill development in electricity and residential electrical systems. By being part of such projects, the device supports the adoption of sustainable consumption habits and empowers communities to actively

contribute to environmental protection, serving as a model for responsible and accessible technological innovation.

References

- [1] M. Poverda, *Eficiencia Energética: Recurso no Aprovechado*. OLADE, 2007. [Online]. Available: <https://upsalesiana.ec/ing33ar9r1>
- [2] O. G. Arellano Bastidas, *Estudio y análisis de eficiencia energética del sistema eléctrico del hospital IESS-Ibarra*. Universidad de las Fuerzas Armadas, Ecuador, 2020. [Online]. Available: <https://upsalesiana.ec/ing33ar9r2>
- [3] A. E. Pilicita-Garrido and D. C. Cevallos-Duque, "Innovación tecnológica de un sistema integral para monitorear el consumo eléctrico," *IngeniUS Revista de Ciencia y Tecnología*, no. 22, pp. 9–16, 2019. [Online]. Available: <https://doi.org/10.17163/ings.n22.2019.01>
- [4] J. I. Peña Paredes and A. P. Trujillo Gavilanes, *Monitorización, análisis y difusión del consumo energético eléctrico de modelo de vivienda del sector residencial de la ciudad de Riobamba*. Escuela Superior Politecnica del Chimborazo, Ecuador, 2014. [Online]. Available: <https://upsalesiana.ec/ing33ar9r4>
- [5] N. Benito Marín, *Optimización de contratación de electricidad para consumidores domésticos y pymes*. Universidad CAros II de Madrid, España, 2016. [Online]. Available: <https://upsalesiana.ec/ing33ar9r5>
- [6] M. Nieves García, *Propuesta y evaluación de tratamientos para la mejora de la eficiencia energética en el sector residencial mediante el desarrollo de experimentos económicos*. Universitat Jaume I, España, 2017. [Online]. Available: <https://upsalesiana.ec/ing33ar9r6>
- [7] A. Costa, F. Ciniello, G. Grzybowski, and R. C. Betini, "Development of a prototype for measurement of electrical energy," *Brazilian Archives of Biology and Technology*, vol. 61, 2018. [Online]. Available: <https://doi.org/10.1590/1678-4324-smart-2018000032>
- [8] C. A. Iza Calapaqui and L. J. Latacunga Pilatasig, *Diseño y construcción de un prototipo para determinar el consumo de energía eléctrica residencial*. Universidad Técnica de Cotopaxi, Ecuador, 2021. [Online]. Available: <https://upsalesiana.ec/ing33ar9r8>

- [9] A. Pitì, G. Verticale, C. Rottondi, A. Capone, and L. Lo Schiavo, “The role of smart meters in enabling real-time energy services for households: The italian case,” *Energies*, vol. 10, no. 2, 2017. [Online]. Available: <https://doi.org/10.3390/en10020199>
- [10] T. AlSkaif, I. Lampropoulos, M. van den Broek, and W. van Sark, “Gamification-based framework for engagement of residential customers in energy applications,” *Energy Research & Social Science*, vol. 44, pp. 187–195, 2018. [Online]. Available: <https://doi.org/10.1016/j.erss.2018.04.043>
- [11] J. López, R. Luna, J. Cervantes, J. Meneeses Ruiz, and J. Hernández, “Aplicación de tecnologías de medición avanzada (ami) como instrumento para reducción de pérdidas,” *Boletín IIE*, vol. 39, pp. 180–191, 12 2015. [Online]. Available: <https://upsalesiana.ec/ing33ar9r11>
- [12] L. M. Estrada Albiño, *Medidor residencial de energía eléctrica con revisión de consumo de forma inalámbrica*. Universidad Tecnica del Norte, Ecuador, 2023. [Online]. Available: <https://upsalesiana.ec/ing33ar9r12>
- [13] D. I. Samaniego Idrovo and D. F. Vele-saca Orellana, *Diseño e implementación de un medidor de energía electrónico para vivienda, con orientación a la prevención de consumo y ahorro energético*. Universidad Politecnica Salesiana, Ecuador, 2016. [Online]. Available: <https://upsalesiana.ec/ing33ar9r13>
- [14] P. C. Vashist and A. Tripathi, “Design and implementation of smart energy meter with real-time pricing,” in *Computational and Experimental Methods in Mechanical Engineering*, V. V. Rao, A. Kumaraswamy, S. Kalra, and A. Saxena, Eds. Singapore: Springer Singapore, 2022, pp. 499–507. [Online]. Available: https://doi.org/10.1007/978-981-16-2857-3_49
- [15] H. K. Patel, T. Mody, and A. Goyal, “Arduino based smart energy meter using gsm,” in *2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU)*, 2019, pp. 1–6. [Online]. Available: <https://doi.org/10.1109/IoT-SIU.2019.8777490>
- [16] E. A. Quintero Salazar, “Medidor electrónico interactivo de consumo de energía eléctrica para uso residencial,” *Prospectiva*, vol. 14, no. 1, pp. 55–70, 2016. [Online]. Available: <https://doi.org/10.15665/rp.v14i1.639>
- [17] A. E. Pilicita-Garrido and D. C. Cevallos-Duque, “Innovación tecnológica de un sistema integral para monitorear el consumo eléctrico,” *Ingenius, Revista de Ciencia y Tecnología*, no. 22, pp. 9–16, 2019. [Online]. Available: <https://doi.org/10.17163/ings.n22.2019.01>
- [18] J. S. Echeverría, , and J. García-Echeverría, “edición avanzada inteligente, retos al consumo responsable del servicio público domiciliario de energía en Colombia,” *Revista chilena de derecho y tecnología*, vol. 11, pp. 47–62, 12 2022. [Online]. Available: <http://dx.doi.org/10.5354/0719-2584.2021.64167>
- [19] A. M. Díaz Rodríguez, J. B. Cogollos Martínez, J. Peña Acción, A. Cogollos Izaguirre, and R. González Álvarez, “Modelo matemático para predecir el consumo de energía eléctrica en la universidad de cienfuegos,” *Universidad y Sociedad*, vol. 12, no. 4, pp. 45–56, 2020. [Online]. Available: <https://upsalesiana.ec/ing33ar9r19>
- [20] Z. de la Cruz Severiche Maury, A. Fernández Vilas, and R. P. Díaz Redondo, “Low-cost hem with arduino and zigbee technologies in the energy sector in colombia,” *Energies*, vol. 15, no. 10, p. 3819, May 2022. [Online]. Available: <http://dx.doi.org/10.3390/en15103819>
- [21] W. Luan, J. Peng, M. Maras, J. Lo, and B. Harapnuk, “Smart meter data analytics for distribution network connectivity verification,” *IEEE Transactions on Smart Grid*, vol. 6, no. 4, pp. 1964–1971, 2015. [Online]. Available: <https://doi.org/10.1109/TSG.2015.2421304>
- [22] T. Hengyu, W. Huanning, H. Yan, Z. Zhihua, Y. Hejun, and L. Yuan, “Real-time monitoring method for smart meter measurement performance,” in *2021 5th International Conference on Smart Grid and Smart Cities (ICSGSC)*, 2021, pp. 72–76. [Online]. Available: <https://doi.org/10.1109/ICSGSC52434.2021.9490438>
- [23] G. Artale, A. Cataliotti, V. Cosentino, S. Guaiana, D. Di Cara, N. Panzavecchia, G. Tiné, N. Dipaola, and M. G. Sambataro, “Pq metrics implementation on low cost smart metering platforms. a case study analysis,” in *2018 IEEE 9th International Workshop on Applied Measurements for Power Systems (AMPS)*, 2018, pp. 1–6. [Online]. Available: <https://doi.org/10.1109/AMPS.2018.8494866>
- [24] A. S. Metering, S. Visalatchi, and K. K. Sandeep, “Smart energy metering and power theft control using arduino & gsm,” in *2017 2nd International Conference for Convergence in Technology (I2CT)*, 2017, pp. 858–961. [Online]. Available: <https://doi.org/10.1109/I2CT.2017.8226251>
- [25] N. S. Živic, O. Ur-Rehman, and C. Ruland, “Evolution of smart metering systems,”

- in *2015 23rd Telecommunications Forum Telfor (TELFOR)*, 2015, pp. 635–638. [Online]. Available: <https://doi.org/10.1109/TELFOR.2015.7377547>
- [26] J. A. Arévalo-López, N. Ávila Peñuela, E. Estupiñan-Escalante, O. A. Parra-Urrego, and D. A. Cano-Tirado, “Design and implementation of an energy metering system to recognize the household electrical energy consumption pattern through an iot network,” in *2021 Joint Conference - 11th International Conference on Energy Efficiency in Domestic Appliances and Lighting & 17th International Symposium on the Science and Technology of Lighting (EEDAL/LS:17)*, 2022, pp. 1–6. [Online]. Available: <https://upsalesiana.ec/ing33ar9r26>
- [27] M. A. Alahmad, P. G. Wheeler, A. Schwer, J. Eiden, and A. Brumbaugh, “A comparative study of three feedback devices for residential real-time energy monitoring,” *IEEE Transactions on Industrial Electronics*, vol. 59, no. 4, pp. 2002–2013, 2012. [Online]. Available: <https://doi.org/10.1109/TIE.2011.2165456>
- [28] OEM. (2023) Openenergymonitor. GitHub. [Online]. Available: <https://upsalesiana.ec/ing33ar9r28>
- [29] ARCERNNR, *Pliego tarifario del servicio público de energía eléctrica año 2023*. Agencia de Regulación y Control de Energía y Recursos Naturales no Renovables, Ecuador, 2023. [Online]. Available: <https://upsalesiana.ec/ing33ar9r29>
- [30] INEN, *NTE INEN-IEC 62053-21: Equipos de medida de energía eléctrica. Requisitos particulares. Parte 21: Contadores estáticos de energía activa (clases 1 y 2)*. Instituto Ecuatoriano de Normalización, Ecuador, 2010. [Online]. Available: <https://upsalesiana.ec/ing33ar9r30>
- [31] ARCERNNR, *Regulación ARCONEL 001/2020*. Agencia de Regulación y Control de Energía y Recursos Naturales no Renovables, Ecuador, 2023. [Online]. Available: <https://upsalesiana.ec/ing33ar9r31>



COMPARATIVE ANALYSIS OF FLOW PATTERNS IN OFF-DESIGN PLANAR AND CONICAL NOZZLE

ANÁLISIS COMPARATIVO DE PATRONES DE FLUJO EN TOBERAS PLANAS Y CÓNICAS FUERA DE DISEÑO

San L. Tolentino^{1,*} , Jorge Mirez¹ 

Received: 27-10-2024, Received after review: 03-12-2024, Accepted: 09-12-2024, Published: 01-01-2025

Abstract

This study aims to analyze the behavior of Mach number and pressure field flow patterns in off-design planar and conical nozzles with a divergent half-angle of 10.85°. Numerical simulations of the flow field were conducted using the ANSYS-Fluent R16.2 software, employing the RANS model and the SAS turbulence model under transient flow conditions. The nozzle pressure ratios (NPR) ranged from 1.97 to 8.91. The results reveal differences in flow patterns, including Mach number and static pressure, between the two nozzle types. Notably, normal shock fronts exhibited varying positions for the same NPR values. The maximum peak flow fluctuation along the centerline of the conical nozzle's divergent section reached Mach 2.844, compared to Mach 2.011 in the planar nozzle, indicating lower flow velocity in the latter. At the nozzle outlet, the flow velocity of the conical nozzle was Mach 2.535, representing a 27.32% increase compared to the planar nozzle, which achieved Mach 1.991. Additionally, the throat area significantly influenced mass flow transit, with the planar nozzle having a larger throat area than the conical nozzle. These findings provide insights into the impact of nozzle geometry on flow characteristics under off-design conditions.

Keywords: Flow fluctuations, RANS model, SAS turbulence model, Shock wave, Flow patterns, Off-design nozzles.

Resumen

En el presente trabajo, el objetivo es determinar el comportamiento de los patrones de flujo del campo de número de Mach y de presión para toberas planas y cónicas fuera de diseño, para el semiángulo de la divergente de 10,85°. Se empleó el código ANSYS-Fluent R16.2 para simular el campo de flujo con el modelo RANS y el modelo de turbulencia SAS para las condiciones de flujo en estado transitorio, para el rango de relaciones de presión de la tobera de NPR 1,97 a 8,91. Los resultados presentan diferentes patrones de flujo de número de Mach y de presión estática entre ambas toberas, donde los frentes de choque normales no tienen las mismas posiciones para un mismo valor de NPR. El pico máximo de la fluctuación del flujo en la línea central de la divergente de la tobera cónica es Mach 2,844, mientras que en la tobera plana es Mach 2,011, por lo que la velocidad del flujo es menor en la tobera plana. La velocidad del flujo a la salida de la tobera cónica es Mach 2,535, la cual es 27,32 % mayor que la velocidad del flujo en la tobera plana, que tiene Mach 1,991. El área de la garganta de la tobera tiene un efecto significativo para el tránsito del flujo másico, ya que el área de la garganta de la tobera plana es mayor con respecto al de la tobera cónica.

Palabras clave: fluctuaciones de flujo, modelo RANS, modelo de turbulencia SAS, onda de choque, patrones de flujo, toberas fuera de diseño

^{1,*}Group of mathematical modeling and numerical simulation, Universidad Nacional de Ingeniería, Lima, Perú.
Corresponding author ✉: stolentino@unexpo.edu.ve.

Suggested citation: Tolentino, S.L. and Mirez, J. "Comparative analysis of flow patterns in off-design planar and conical nozzle," *Ingenius, Revista de Ciencia y Tecnología*, N.º 33, pp. 115-127, 2025, DOI: <https://doi.org/10.17163/ings.n33.2025.10>.

1. Introduction

The geometries of convergent-divergent nozzles used in the aerospace field significantly influence the development of the flow field. Consequently, the geometry of the nozzle's inner wall and the resulting flow turbulence remain recurring topics of interest, particularly under overexpanded, adapted, and under expanded viscous flow conditions [1]. Since the advent of technological advancements in supersonic nozzles, their geometries have undergone continuous evolution to optimize and regulate flow transit. Among the prominent nozzle geometries are conical nozzles, contoured or bell-shaped designs (both full-length and shortened), plug or aerospike nozzles (full-length and truncated), expansion-deflection nozzles, planar nozzles, and others [1–6]. Research in this area has extended to the method of characteristics (MOC) [7, 8], as well as the development of approximate mathematical models to simulate flow using computational fluid dynamics (CFD) tools [9, 10]. Furthermore, studies have explored the solution of mathematical models for isentropic flow involving analytical equations that cannot be inverted through algebraic procedures [11–14].

Supersonic conical nozzles with optimal performance typically have a divergent half-angle α ranging from 12° to 18° . When $\alpha < 12^\circ$, these nozzles are classified as off-design [1]. The same principle applies to planar nozzles with no curvature in the divergent section.

Under overexpanded flow conditions, various shock wave phenomena occur in the divergent section of the nozzle, including normal, oblique, reflected, and internal shock waves. Figure 1 illustrates the shock wave structure in the divergent section of a supersonic nozzle [15], along with an image of a shock wave captured using the Schlieren technique, as reported by Hunter [6], [16].

The central flow region contains the shock front, flanked by oblique and reflected waves that interact with pressure and temperature gradients [6], [17, 18]. In the flow region adjacent to the nozzle walls, lateral pressure loads are observed before and after the shock. The boundary layer interacts with the shock waves, and downstream of the flow separation point, backpressure and flow recirculation occur, accompanied by the formation of vortices. The effects of temperature and friction are prominent in the flow region near the walls, where free shock separation (FSS) and restricted shock separation (RSS) phenomena are observed [19–22]. At the nozzle's edges, such as the exit edge, Prandtl-Meyer expansion waves occur [17]. The supersonic jet flow discharged into the atmosphere forms a plume containing shock wave structures, which are influenced by the nozzle pressure ratio (NPR) [19].

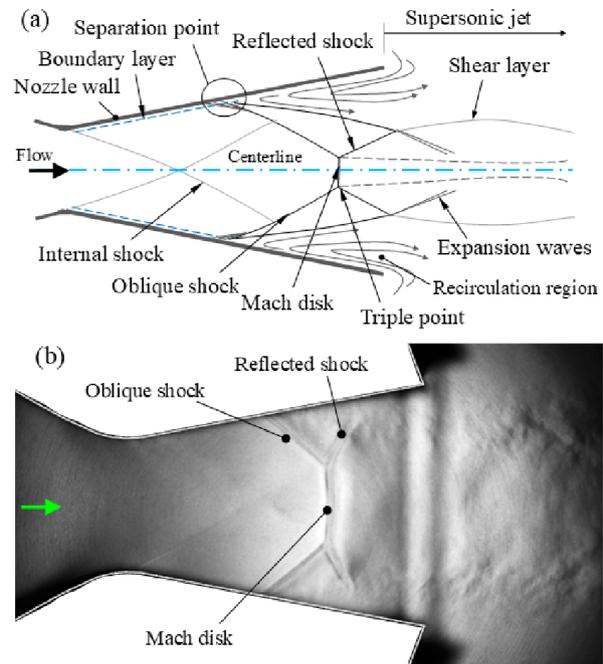


Figure 1. (a) Shock wave structure for the overexpanded flow condition [15]. (b) Experimental image of the shock wave structure in a planar nozzle, as reported by Hunter [6]

Several significant studies have explored the behavior of flow in off-design nozzles. Hunter [6] carried out experimental tests on flow separation in a planar nozzle with a divergent angle $\alpha = 11,01^\circ$, revealing that overexpanded flow was dominated by shock-induced boundary layer separation. Likewise, Verma and Mani Sankar [18] investigated planar nozzles with divergent angles $\alpha = 5,7^\circ$, $\alpha = 7,5^\circ$ and $\alpha = 10,7^\circ$, identifying asymmetry in normal, oblique, and reflected shock wave fronts, as well as non-symmetric boundary layer separation between the top and bottom walls. Tolentino et al. [23] conducted computational simulations for planar nozzles with $\alpha = 9^\circ$, $\alpha = 11,01^\circ$ and $\alpha = 13^\circ$, reporting that increasing the divergent angle led to increased flow velocity in the central region at the nozzle outlet. Arora and Vaidyanathan [24] investigated planar nozzles with double divergence, finding that the angle of inflection significantly influenced the shock structure, with additional expansion occurring at the second divergence.

Tolentino et al. [25] simulated the flow field in planar nozzles with a straight cutthroat for $\alpha = 11,01^\circ$. They found that as throat length increased, the shock train evolved within the throat section, affecting flow development in the divergent.

Mason et al. [26] experimentally studied planar nozzles with divergent angles $\alpha = 1,21^\circ$, $\alpha = 10,85^\circ$ and $\alpha = 11,24^\circ$. Their results showed that the convergent and throat contour significantly affected flow behavior. For $\alpha = 1,21^\circ$, pressure fluctuations were observed along the divergent wall due to the presence of a shock train. In contrast, for $\alpha = 10,85^\circ$, only a pressure jump

at the flow separation point was observed due to the normal shock front.

Tolentino et al. [15] studied conical nozzles with a straight-cut throat for $\alpha = 10^\circ$, showing that increasing throat length led to the evolution of the shock train and a reduction in thrust at the nozzle exit. Tolentino and Mirez [27] further analyzed flow patterns in conical nozzles, revealing that an optimal throat length prevents the formation of a shock train.

Other experimental studies have also investigated flow behavior in nozzles. Wagner and Schleichriem [28] examined flow turbulence in planar nozzles with wall curvature, reporting non-symmetric shock wave propagation. Bourgoing and Reijasse [29] demonstrated that wall roughness affects flow development, resulting in varying shock wave asymmetry configurations. Faheem et al. [30] conducted experiments with nozzles emitting multiple supersonic jets, finding that as the number of jets increased, the scattering rate decreased due to reduced drag, with notable differences in the cores of the supersonic jets.

Computational fluid dynamics (CFD) simulations [9] have proven effective in reproducing flow turbulence and determining thermodynamic parameters in regions where experimental imaging is not feasible. Recent research has focused on modeling turbulence for Mason et al.'s [26] experimental flat nozzle geometry, with extensions to conical nozzle geometries.

This study aims to analyze the behavior of flow patterns in planar and conical nozzles classified as off-design nozzles. Both nozzle types share identical geometric dimensions projected in a 2D computational domain. Section 2 outlines the methodology, including a description of the experimental nozzle geometry used by Mason et al. [26] to record pressure readings at different nozzle pressure ratios (NPR) under cold flow

conditions. The mathematical foundation and computational solution methods are also detailed, along with numerical convergence analysis and validation of the SAS turbulence model employed in simulations. Section 3 presents the numerical results, focusing on Mach number flow fields and static pressure distributions. Finally, Section 4 provides the conclusions of the analysis, summarizing the findings and their implications for future research.

2. Materials and method

2.1. Experimental nozzle

The experimental planar nozzle (Model B1) under study, as illustrated in Figure 2, was previously utilized by Mason et al. [26]. Pressure experiments for cold flow conditions were conducted at the NASA Langley Research Center's 16-Foot Transonic Tunnel Complex. These experimental tests covered a nozzle pressure ratio (NPR) range from 1.97 to 8.91. The planar nozzle [26] was designed for a Mach number of 2.08, an NPR of 8.81, an area ratio $A_e/A_t = 1,8$, and a divergent half-angle $\alpha = 10,85^\circ$, classifying it as an off-design nozzle since $\alpha < 12^\circ$. A_e represents the nozzle outlet area, and A_t denotes the throat area. The width of the planar nozzle is 10.157 cm.

It is noteworthy that the geometric dimensions of the planar nozzle, when projected onto a 2D plane (Figure 2a), are applied to the conical nozzle geometry. For the conical nozzle, the radii are defined as h_i for the inlet, h_t for the throat and h_e for the outlet, with a divergent half-angle $\alpha = 10,85^\circ$. The design conditions for the conical nozzle include a Mach number of 2.713, an NPR of 23.768, and an area ratio A_e/A_t of 3,224.

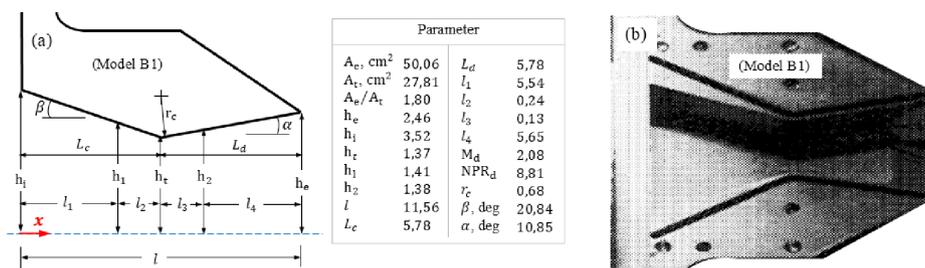


Figure 2. (a) Schematic representation of the planar nozzle, illustrating geometric parameters (units in centimeters). (b) Experimental planar nozzle used by Mason et al. [26] to measure the flow pressure along the inner walls

2.2. Computational domain

The spatial projection of the planar and conical nozzles was considered in 2D computational domains due to their symmetry. These domains were constructed using the geometric data reported by Mason et al. [26], as

depicted in Figure 2. Figure 3 illustrates the computational domain along with the meshed domain for Grid 4, which consists of 30,736 quadrilateral cells. Grid 4 represents a refined mesh with improved cell distribution, previously evaluated for performance compared to other meshed domains. The geometric dimensions are

parameterized, with the centerline aligned along the x -direction, and the divergent section denoted as L_d . The convergent section of the nozzle spans $0 \leq x/L_d \leq 1$, the divergent section spans $1 \leq x/L_d \leq 2$, and the atmosphere section extends from $2 \leq x/L_d \leq 8$. This computational domain setup is suitable for capturing the flow turbulence and simulating the formation

of the plume in the supersonic jet as it discharges into the atmosphere. It is important to note that 2D computational domains are appropriate for symmetric geometries, offering a significant reduction in iterative computational time and resource usage compared to 3D computational domains.

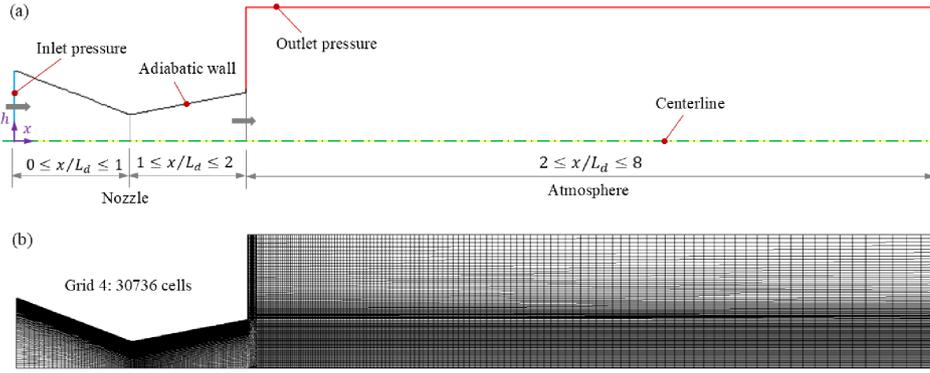


Figure 3. (a). 2D computational domain illustrating the applied boundary conditions (b) Structured mesh for Grid 4, consisting of 30736 quadrilateral cells

The boundary conditions applied to the 2D computational domain for both the planar and conical nozzles are as follows: At the nozzle inlet, stagnation pressure P_o data are specified for eight cases corresponding to nozzle pressure ratios (NPR) of 1.97, 2.94, 3.92, 4.88, 5.84, 6.81, 7.79, and 8.91, where $NPR = P_o/P$ and the pressure of the local atmosphere $P = 101,3$ kPa. The total temperature T_o is set to 300 K for all eight cases. At the nozzle outlet, the local atmospheric pressure and temperature are uniformly set to $P = 101,3$ kPa and $T = 300$ K for all cases. It is important to note that the pressure and temperature data were obtained from the work of Mason et al. [26].

The nozzle walls are modeled as adiabatic, meaning no heat transfer occurs through them. The mesh is refined in the regions adjacent to the walls to account for the presence of flow shear stresses. At the walls, the flow velocity is zero due to the no-slip condition. Along the centerline, the flow velocity is zero in the perpendicular direction for the planar nozzle and in the radial direction for the conical nozzle. The effects of gravity on flow turbulence are neglected, as the computational domains are two-dimensional and represent symmetric geometries.

The airflow is treated as an ideal gas with the following thermodynamic properties: gas constant $R = 287$ J/(kg·K), specific heat ratio $k = 1,4$, specific heat at constant pressure $C_p = 1006,43$ J/(kg·K) and thermal conductivity $k_t = 0,042$ W/(m·K) [31].

2.3. Mathematical fundamentals

The turbulence of the viscous flow field in a transient state is simulated using the ANSYS-Fluent R16.2 software [31], which employs the finite volume method (FVM) [9]. The Reynolds-Averaged Navier-Stokes (RANS) equations are utilized, as they provide an efficient and appropriate framework for obtaining approximate solutions to compressible flow turbulence. The governing equations for the calculation of compressible flow include the conservation of mass (Equation (1)), momentum (Equation (2)), and energy (Equation (3)) [9, 10], [31]. These equations, expressed in compact form, are as follows:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho u_i) = 0 \quad (1)$$

$$\frac{\partial}{\partial t} (\rho u_i) + \nabla \cdot (\rho u_i u_j) = -\nabla p + \nabla \cdot (\bar{\tau}) + \nabla \cdot (-\rho \overline{u'_i u'_j}) \quad (2)$$

$$\frac{\partial}{\partial t} (\rho E) + \nabla \cdot (u_i (\rho E + p)) = \nabla \cdot (k_{eff} \nabla T + (\bar{\tau}_{eff} \cdot u_i)) \quad (3)$$

Where t is the time, ρ is the density, u is velocity, p is the pressure and $\bar{\tau}$ is the stress tensor. The term $-\rho \overline{u'_i u'_j}$ represents the Reynolds stress, where turbulence models are applied to close Equation (2). Additionally, E denotes the total energy, k_{eff} is the effective thermal conductivity, T is the temperature, and $\bar{\tau}_{eff}$ is the effective tensor.

The flow is modeled as an ideal gas; hence, the ideal gas equation of state (Equation (4)) is applied. The equations for pressure ratio (Equation (5)) and temperature (Equation (6)) are expressed as functions of the Mach number:

$$p = \rho RT \quad (4)$$

$$\frac{P_o}{P} = \left(1 + \frac{k-1}{2} M^2\right)^{\frac{k}{k-1}} \quad (5)$$

$$\frac{T_o}{T} = 1 + \frac{k-1}{2} M^2 \quad (6)$$

Where R is the universal ideal gas constant, P_o is the stagnation pressure, P is the static pressure, T_o is the stagnation temperature, T is the static temperature, and k is the ratio of specific heats. The Mach number, denoted as M , is categorized as follows: subsonic flow $0,3 \leq M \leq 0,8$, transonic flow $0,8 \leq M \leq 1,2$, supersonic flow $1,2 \leq M \leq 5$, hypersonic flow $M > 5$, and sonic flow $M = 1$. For incompressible flow, $M < 0,3$ is considered [32].

The Scale-Adaptive Simulation (SAS) model [33] was employed to simulate the turbulence of viscous flow in a transient state. Additionally, Sutherland's law [10], [32] was applied to model the flow viscosity as a function of temperature. Sutherland's law, derived from an approximation of the kinetic theory of gases, accounts for the idealization of intermolecular force potentials and is expressed as equation (7):

$$\frac{\mu}{\mu_o} = \left(\frac{T}{T_o}\right)^{\frac{2}{3}} \frac{T_o + S}{T + S} \quad (7)$$

Where μ is the flow viscosity, $\mu_o = 1,716$ kg/m·s is the reference viscosity, $T_o = 273,11$ K is the reference temperature, and $S = 110,56$ K is the effective temperature. The effective temperature S , also known as the Sutherland constant, is tabulated for various temperature ranges depending on the type of gas.

2.4. Computational solution method

The simulations in ANSYS-Fluent R16.2 were conducted using a density-based solver with a transient

formulation. The 2D computational domain was planar for the planar nozzle and axisymmetric for the conical nozzle. An implicit formulation was employed, with Roe-FDS selected as the flux type. Spatial discretization was performed using a least squares cell-based approach and second-order upwind schemes, while the transient formulation utilized a second-order implicit method. Hybrid initialization was applied, and a residual convergence criterion of 1×10^{-5} was set. The flow field simulations for the nozzle pressure ratio (NPR) range of 1.97–8.91 were resolved in 7,200 to 21,000 iterations.

2.5. Numerical convergence analysis

The computational domains were meshed using the ANSYS-Meshing platform. Four quadrilateral cell grids with varying densities were generated: grid 1 with 19,249 cells, grid 2 with 25,658 cells, grid 3 with 27,311 cells, and grid 4 with 30,736 cells. The Mach number flow field was simulated using the Scale-Adaptive Simulation (SAS) turbulence model [33] for an NPR of 8.91. Among the four grids, grid 4 exhibited the lowest Y -plus (y^+) values, particularly in the shear stress distribution evaluated along the planar nozzle wall (Figure 4a). Regarding shear stress values, the curves for all grids are generally consistent and superimposed, except in the region near the nozzle outlet (Figure 4b).

For the domain meshed with grid 4, the SAS turbulence model [33] was compared with the DES SA [34], DES SST $k-w$ [35], and DES $k-\epsilon$ [36] turbulence models, as illustrated in Figure 4c. The results indicate that the average Mach number at the nozzle exit is similar for the SAS and DES SA turbulence models, as shown in Table 1. The SAS turbulence model was selected for simulating the flow field due to its slight advantage in reducing computational time during iterative calculations. Grid 4 (Figure 3b) features a structured grid composed of quadrilateral cells, with refined regions in the flow zones adjacent to the convergent section ($0 \leq x/L_d \leq 1$) and divergent section ($1 \leq x/L_d \leq 2$), as well as along the walls. Additionally, the grid includes a progressively distributed cell structure in the atmosphere section ($2 \leq x/L_d \leq 8$).

Table 1. Average Mach number at the exit of the planar nozzle for NPR 8.91, evaluated using four different turbulence models

Turbulence model:	SAS [33]	DES SA [34]	DES SST $k-w$ [35]	DES $k-\epsilon$ [36]
Average Mach number:	2.066	2.066	2.065	2.064
Error (%):	0.816	0.816	0.864	0.912

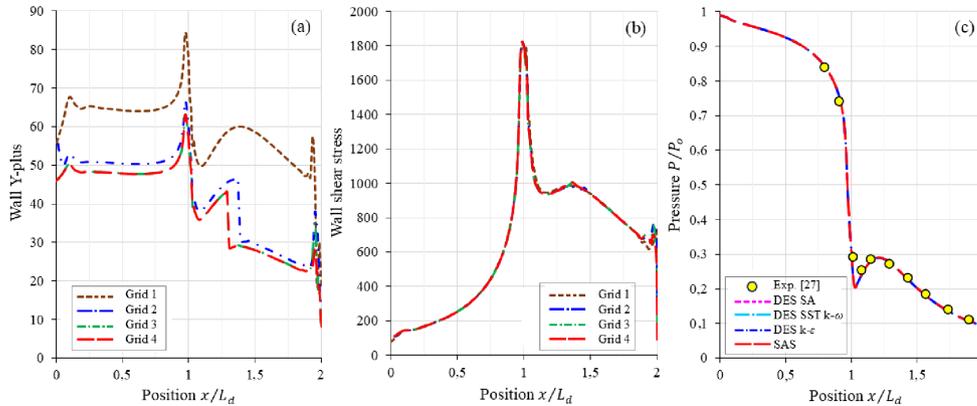


Figure 4. Profiles of curves evaluated along the wall of the planar nozzle: (a) Wall Y-plus, (b) Wall shear stress, and (c) Pressure. Position: convergent section ($0 \leq x/L_d \leq 1$) and divergent section ($1 \leq x/L_d \leq 2$)

3. Results and discussion

This section presents the numerical results for the Mach number flow field (Figure 5) and static pressure distribution (Figure 6) across the NPR range of 1.91 to 8.91. In the flow field visualizations, regions with smaller magnitudes are represented in blue, while regions with larger magnitudes are represented in red. For the planar nozzle, the flow is overexpanded within the NPR range of 1.91 to 7.79 and becomes tailored at NPR 8.91. In contrast, for the conical nozzle, the flow remains overexpanded throughout the entire NPR range of 1.91 to 8.91.

For the planar nozzle, the Mach number flow field (Figure 5a) and static pressure field (Figure 6a) indicate that, at NPR 1.97 and NPR 2.94, the normal shock wave front is located inside the nozzle. For NPR values at or above 3.92, the shock front moves outside the nozzle. In the case of the conical nozzle, the Mach number (Figure 5b) and static pressure (Figure 6b) flow fields reveal that, within the NPR range of 1.97 to 4.88, the normal shock wave front remains inside the nozzle. For NPR values equal to or greater than 5.84, the shock front shifts outside the nozzle.

The evolution of the Mach number flow field and static pressure field for both the planar and conical nozzles demonstrates that the flow regime exhibits distinct behaviors at the same NPR value. As the NPR increases, the shock wave structure evolves, and the flow separation point shifts closer to the nozzle exit.

In the planar nozzle, the distribution of internal shocks is evident both within the nozzle and in the supersonic jet discharging into the atmosphere. However, for the same NPR magnitude, the distribution of internal shocks differs in the conical nozzle. The intensity of shock front displacement is greater in the planar nozzle compared to the conical nozzle.

The trajectories of the centerline profiles for the planar and conical nozzles are illustrated in Figure 7 for Mach number and Figure 8 for static pressure.

In the divergent section ($1 \leq x/L_d \leq 2$) of the planar nozzle (Figure 7a and Figure 8a), velocity and static pressure fluctuations are observed for the NPR range of 2.94 to 8.91. The maximum peak flow velocity occurs at $x/L_d = 1,486$, where the Mach number reaches 2.011, accompanied by a pressure drop of $P/P_o = 0,125$ (Table 2). At $x/L_d = 1,626$, the flow decelerates to Mach 1.819 with a corresponding pressure drop of $P/P_o = 0,169$. At the nozzle outlet, $x/L_d = 2$, the flow velocity reaches Mach 1.991 with a pressure drop of $P/P_o = 0.129$.

In contrast, within the divergent section ($1 \leq x/L_d \leq 2$) of the conical nozzle (Figure 7b and Figure 8b), for the NPR range of 3,92 to 8,91, the peak flow velocity occurs at $x/L_d = 1,577$, reaching Mach 2.844 with a corresponding pressure drop of $P/P_o = 0,034$ (Table 2). The position of this maximum velocity peak is shifted 6.12% downstream compared to the peak position at $x/L_d = 1,486$ in the planar nozzle (Figure 7a). At $x/L_d = 1,633$, the flow decelerates to Mach 2.024 with a pressure drop of $P/P_o = 0.122$, exhibiting a positional shift of 0.43%.

At the outlet of the conical nozzle (Figure 7b and Figure 8b), at the same $x/L_d = 2$ position as the planar nozzle (Figure reffig7a and Figure reffig8a), the flow reaches Mach 2.535 with a corresponding pressure drop of $P/P_o = 0.054$. This represents a 27.32% increase in velocity and a 58.13% decrease in pressure compared to the planar nozzle. It is observed that at the outlet of the conical nozzle, a higher velocity is achieved along the centerline of the flow's central region. However, the overall mass flow rate is lower. This difference is attributed to the throat area of the conical nozzle $A_t = 5,896 \text{ cm}^2$, which restricts the passage of mass flow at sonic velocity. This throat area is 78.79% smaller than that of the planar nozzle $A_t = 27,81 \text{ cm}^2$.

In the atmospheric region ($2 \leq x/L_d \leq 8$) the flow discharge exhibits notable fluctuations. For the planar nozzle (Figures 7a and 8a), at NPR 8.91, the velocity

reaches a maximum peak of Mach 2.522 with a pressure drop of $P/P_o = 0,056$ at $x/L_d = 2,95$. Conversely, the conical nozzle (Figures 7b and 8b) achieves a peak velocity of Mach 2.976 with a pressure drop of $P/P_o = 0,027$ at $x/L_d = 2,408$, reflecting an approximately

18% increase in velocity at this position. Further downstream, the planar nozzle exhibits milder velocity and pressure fluctuations, while these fluctuations are more pronounced in the conical nozzle.

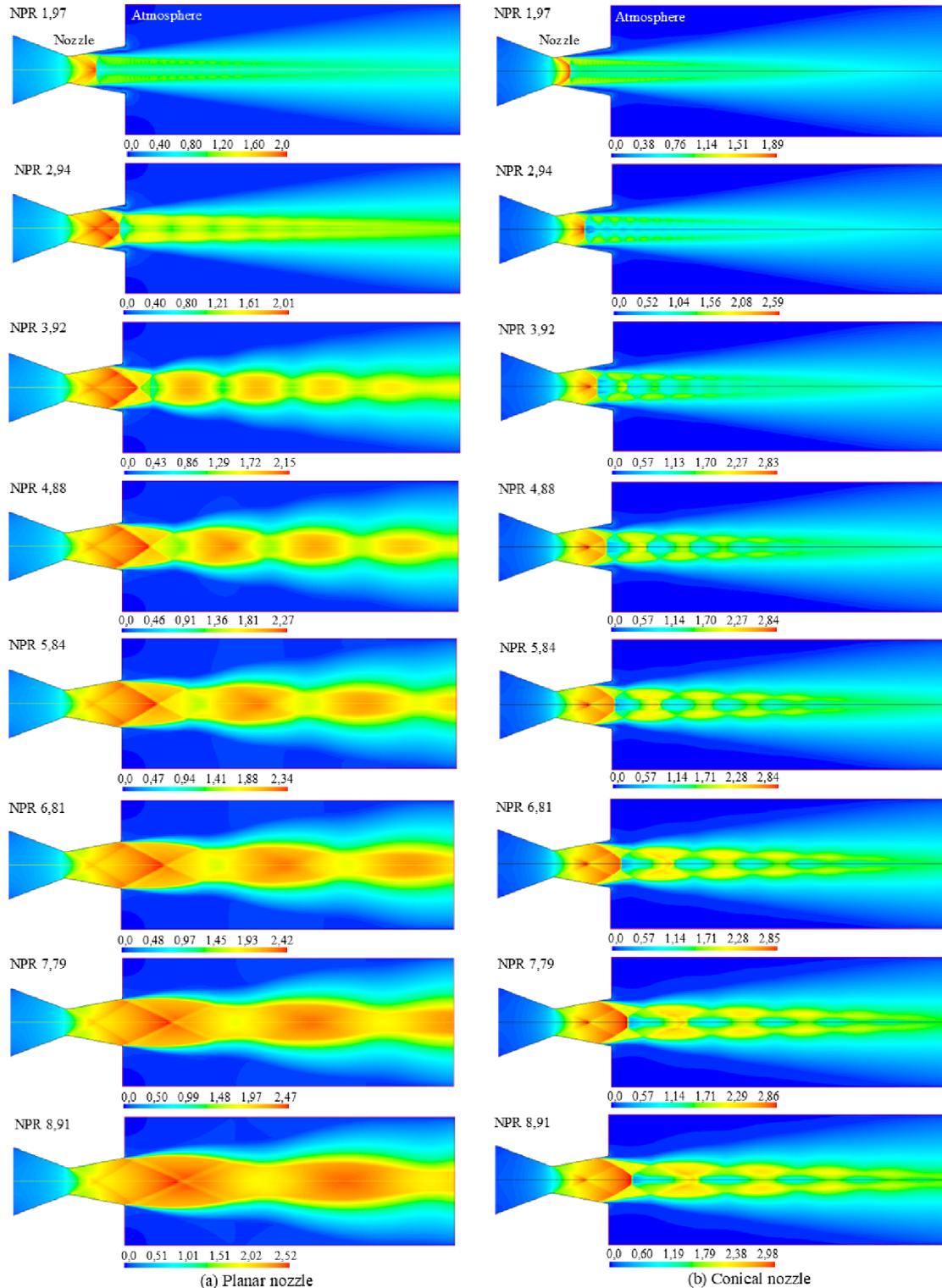


Figure 5. Mach number flow field patterns. Range: NPR 1,97 to NPR 8,91. (a) Planar nozzle. (b) Conical nozzle

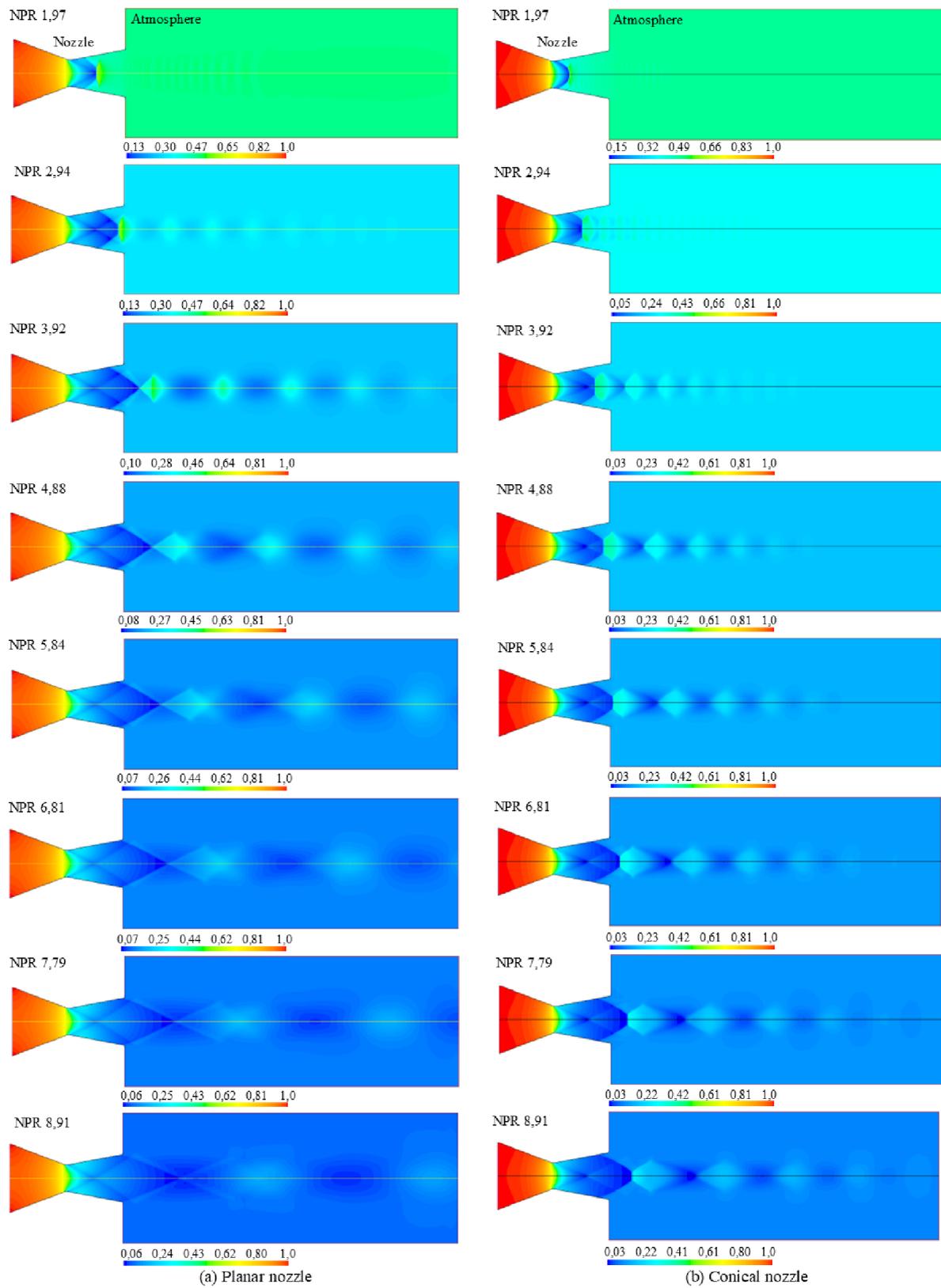


Figure 6. Pressure flow field patterns. Range: NPR 1,97 to NPR 8,91. (a) Planar nozzle. (b) Conical nozzle

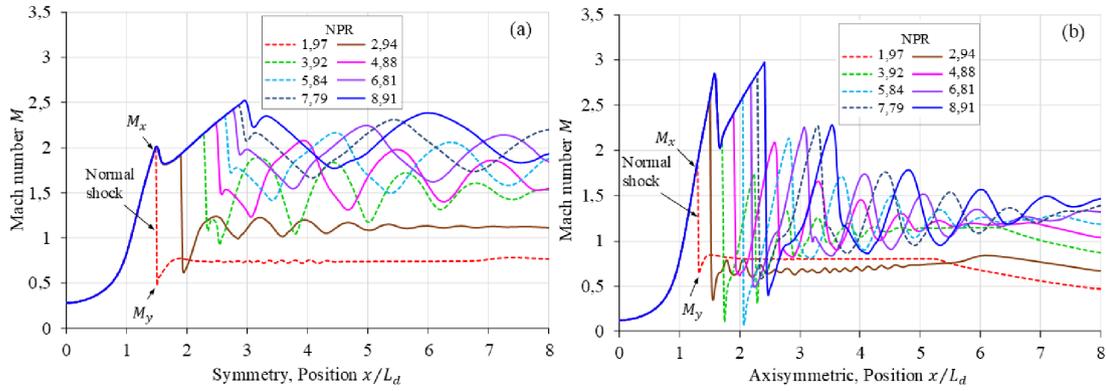


Figure 7. Mach number profile patterns evaluated along the centerline. (a) Planar nozzle. (b) Conical nozzle. Position: convergent section ($0 \leq x/L_d \leq 1$), divergent section ($1 \leq x/L_d \leq 2$) and atmosphere section ($2 \leq x/L_d \leq 8$)

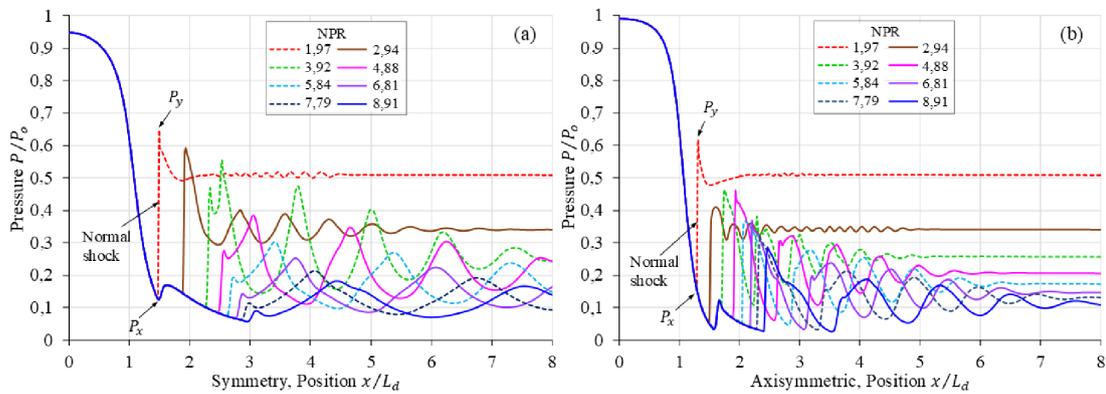


Figure 8. Pressure profile patterns evaluated along the centerline. (a) Planar nozzle. (b) Conical nozzle. Position: convergent section ($0 \leq x/L_d \leq 1$), divergent section ($1 \leq x/L_d \leq 2$) and atmosphere section ($2 \leq x/L_d \leq 8$)

Table 2. Position of the maximum and minimum fluctuations in the central region of the supersonic flow along the centerline of the nozzle’s divergent section

Planar nozzle. Range: NPR 2,94 to NPR 8,91		
Position x/L_d	Mach number M	Position P/P_0
1,486	2,011	0.125
1,626	1,819	0.169
2	1,991	0.129
Conical nozzle. Range: NPR 3,92 to NPR 8,91		
Position x/L_d	Mach number M	Position P/P_0
1,577	2,844	0.034
1,633	2,024	0.122
2	2,535	0.054

Table 3 presents the positions of the normal shock front evaluated along the centerline of the planar nozzle (Figure ??a) and the conical nozzle (Figure ??b). M_x represents the maximum flow velocity at the start of the shock front, while M_y corresponds to the minimum flow velocity at its end. Figures 7 and 8 also show the minimum pressure P_x and the maximum pressure P_y of the shock front for reference.

In the divergent section of the planar nozzle, two

normal shock fronts are observed for NPR 1.97 and NPR 2.94. At the start of the shock front, the flow achieves supersonic velocities ranging from Mach 1.995 to 1.925. By the end of the shock front, the flow transitions to subsonic velocities in the range of Mach 0.473 to 0.610. Outside the nozzle, six normal shock fronts are identified for NPR values of 3.92, 4.88, 5.84, 6.81, 7.79, and 8.91. At the start of these shock fronts, the flow exhibits supersonic velocities ranging from Mach 2.147 to 2.522. At the end of the shock fronts, the flow transitions to transonic and supersonic velocities within the range of Mach 1.072 to 2.221.

In the divergent section of the conical nozzle, four shock fronts are observed for NPR values of 1.97, 2.94, 3.92, and 4.88. At the start of the shock fronts, the flow reaches supersonic velocities ranging from Mach 1.877 to 2.406. By the end of the shock fronts, the flow transitions to subsonic velocities in the range of Mach 0.634 to 0.615. Outside the nozzle, four additional shock fronts are present for NPR values of 5.84, 6.81, 7.79, and 8.91. At the start of these shock fronts, the flow remains supersonic, with velocities ranging from Mach 2.589 to 2.976. At the end of the shock fronts,

the flow transitions to subsonic velocities within the range of Mach 0.072 to 0.4.

It is observed that, at the beginning of the shock fronts for NPR values equal to or greater than 2.94, the flow in the conical nozzle achieves higher supersonic velocities compared to the planar nozzle. At the end of the shock front, the flow in the conical nozzle transitions to subsonic velocity. In contrast, the flow in the planar nozzle exhibits a broader range of behaviors at the end of the shock front, including subsonic, transonic, and supersonic velocities.

The offsets in the positions of the normal shock fronts in the conical nozzle are summarized in Table 3. For the same NPR value, the shock front in the conical nozzle is shifted further to the extreme left compared to the shock front in the planar nozzle. The smallest displacement of the shock front position occurs at NPR 1.97, with the shock onset offset by 12.47% and the shock end by 11.95%. Conversely, the largest displacement is observed at NPR 3.92, with the shock onset offset by 24.44% and the shock end by 25.38%.

Table 3. Positions of normal shock fronts evaluated along the centerline in the divergent section ($1 \leq x/L_d \leq 2$) and atmosphere section ($2 \leq x/L_d \leq 8$)

NPR	Nozzle planar				Conical nozzle			
	Position x/L_d	Mach number M_x	Position x/L_d	Mach number M_y	Position x/L_d	Mach number M_x	Position x/L_d	Mach number M_y
1.97	1.475	1.995	1.497	0.473	1.291	1.877	1.318	0.634
2.94	1.888	1.925	1.935	0.61	1.497	2.587	1.554	0.342
3.92	2.254	2.147	2.336	1.072	1.703	2.128	1.743	0.103
4.88	2.466	2.267	2.57	1.476	1.888	2.406	2	0.615
5.84	2.615	2.344	2.734	1.721	2.045	2.589	2.061	0.072
6.81	2.759	2.415	2.865	1.924	2.179	2.739	2.239	0.486
7.79	2.837	2.466	2.979	2.067	2.286	2.852	2.354	0.579
8.91	2.95	2.522	3.103	2.221	2.408	2.976	2.466	0.4

Table 4. Leftward offset of normal shock front positions in the conical nozzle relative to the planar nozzle

	NPR							
	1.97	2.94	3.92	4.88	5.84	6.81	7.79	8.91
Normal shock front:								
position offset (x/L_d) in percentages (%)								
Start of the shock, M_x:	12.47	20.7	24.44	23.48	21.79	21.02	19.42	18.37
End of the shock, M_y:	11.95	19.68	25.38	22.17	24.61	21.84	20.98	20.52

Similar findings regarding the overexpanded flow and divergent shock structures in planar nozzles have been reported in [6], [23], and in conical nozzles in [15], [27], where the oblique, reflected, and normal shock fronts are influenced by wall geometry and divergence angle. These studies demonstrate that the geometry of supersonic nozzles significantly impacts the development of the flow regime. For the same geometric dimensions projected in the plane, as shown in Figure 2, the planar nozzle and conical nozzle exhibit distinct cross-sectional areas at both the throat and the nozzle outlet [26]. The cross-sectional area of the planar nozzle is larger than that of the conical nozzle, resulting in a smaller mass flow through the throat of the conical nozzle. It is important to note that numerical results are influenced by the simplifications of the mathematical models, the density of cells in the computational mesh, and the turbulence models employed. Consequently, validation with experimental data was essential. The SAS turbulence model, used in the sim-

ulations, was validated to ensure that the numerical results closely approximate the physical phenomena under investigation.

4. Conclusions

Based on the results obtained from computational simulations of the flow field in planar and conical nozzles with a divergence half-angle $\alpha = 10,85^\circ$, which are classified as off-design nozzles, several key observations can be made:

The behavior of the Mach number and static pressure flow fields differs notably between the two nozzle types for the same NPR value. Within the NPR range of 1.97 to 8.91, the shock wave structures evolve distinctly, with their positions and flow separation points progressively shifting toward the nozzle outlet as NPR increases.

In the divergent section of the planar nozzle, the maximum peak flow velocity is observed at $x/L_d =$

1,486, reaching Mach 2.011 with a pressure drop of $P/P_o = 0,125$. In contrast, the conical nozzle achieves its maximum flow velocity at $x/L_d = 1,577$, with Mach 2.844 and a corresponding pressure drop of $P/P_o = 0,034$. This position is shifted 6.12% downstream compared to the $x/L_d = 1,486$ position in the planar nozzle.

In the central region of the flow, the velocity at the nozzle outlet differs significantly between the two geometries. The planar nozzle achieves a flow velocity of Mach 1.991, whereas the conical nozzle reaches Mach 2.535, representing a velocity increase of 27.32% relative to the planar nozzle.

In the atmospheric region at NPR 8.91, where the supersonic jet plume forms, the maximum flow velocity for the planar nozzle is Mach 2.522. In comparison, the conical nozzle achieves Mach 2.976, reflecting an 18% increase in flow velocity. These findings highlight the critical role of the throat area in determining mass flow transit. The larger throat area of the planar nozzle facilitates greater mass flow, while the smaller throat area of the conical nozzle restricts flow, leading to the observed differences in velocity and pressure profiles. This underscores the significant impact of nozzle geometry on flow behavior, particularly in off-design configurations.

Funding

This research was supported by the Vice-Rectorado de Investigación (VRI) of the Universidad Nacional de Ingeniería (UNI), Lima, Peru.

References

- [1] G. P. Sutton and O. Biblarz, *Rocket Propulsion Elements*. John Wiley & Sons, 2016. [Online]. Available: <https://upsalesiana.ec/ing33ar10r1>
- [2] G. Scarlatella, M. Tajmar, and C. Bach, “Advanced nozzle concepts in retro-propulsion applications for reusable launch vehicle recovery: a case study,” in *72nd International Astronautical Congress (IAC), Dubai, United Arab Emirates*, 2021. [Online]. Available: <https://upsalesiana.ec/ing33ar10r2>
- [3] G. Hagemann, A. Preuss, J. Kretschmer, F. Grauer, M. Frey, R. Ryden, and R. Stark, *Technology Investigation for High Area Ratio Nozzle Extensions*. 39th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit. [Online]. Available: <https://doi.org/10.2514/6.2003-4912>
- [4] F. J. Malina, “Characteristics of the rocket motor unit based on the theory of perfect gases,” *Journal of the Franklin Institute*, vol. 230, no. 4, pp. 433–454, 1940. [Online]. Available: [https://doi.org/10.1016/S0016-0032\(40\)91348-5](https://doi.org/10.1016/S0016-0032(40)91348-5)
- [5] S. Zivkovic, m. Milinovic, and N. Adamec, “Eksperimentalno i numericko istrazivanje super-sonicnog ravanskog mlaznika sa vektorisanim potiskom mehanickim preprekama,” *FME Transactions*, vol. 42, no. 3, pp. 205–211, 2014. [Online]. Available: <https://doi.org/10.5937/fmet1403205Z>
- [6] C. A. Hunter, “Experimental investigation of separated nozzle flows,” *Journal of Propulsion and Power*, vol. 20, no. 3, pp. 527–532, 2004. [Online]. Available: <https://doi.org/10.2514/1.4612>
- [7] J. D. Anderson, *Hypersonic and High-Temperature Gas Dynamics, Third Edition*. AIAA Education Series, 2019. [Online]. Available: <https://doi.org/10.2514/5.9781624105142.0000.0000>
- [8] K. Q. Zaman and A. F. Fagan, *Flow, noise and thrust of supersonic plug nozzles*. AIAA SCITECH, 2024. [Online]. Available: <https://doi.org/10.2514/6.2024-2305>
- [9] J. H. Ferziger and M. Peric, *Computational Methods for Fluid Dynamics*. Springer Berlin, Heidelberg, 2012. [Online]. Available: <https://doi.org/10.1007/978-3-642-56026-2>
- [10] H. Schlichting and K. Gersten, *Boundary-Layer Theory*. Springer Berlin, Heidelberg, 2016. [Online]. Available: <https://doi.org/10.1007/978-3-662-52919-5>
- [11] S. L. Tolentino, “Comparative analysis of 2d simulations and isentropic equations for compressible flow in experimental nozzles,” *INCAS BULLETIN*, vol. 15, pp. 111–125, 09 2023. [Online]. Available: <https://doi.org/10.13111/2066-8201.2023.15.3.9>
- [12] —, “Empirical equation of the mach number as a function of the stagnation pressure ratio for a quasi- one-dimensional compressible flow,” *FME Transactions*, vol. 51, 03 2023. [Online]. Available: <http://dx.doi.org/10.5937/fme2302149T>
- [13] J. Majdalani and B. Maicke, “Explicit inversion of stodola’s area-mach number equation,” *Journal of Heat Transfer*, vol. 133, p. 071702, 07 2011. [Online]. Available: <http://dx.doi.org/10.1115/1.4002596>
- [14] A. Ferrari, “Exact solutions for quasi-one-dimensional compressible viscous flows in conical nozzles,” *Journal of Fluid Mechanics*, vol. 915, p. A1, 2021. [Online]. Available: <https://doi.org/10.1017/jfm.2020.1158>

- [15] S. L. Tolentino, J. Mírez Tarrillo, and S. Caraballo F., “Numerical analysis of the shock train in conical nozzles with straight-cut throats,” *FME Transactions*, vol. 52, pp. 186–195, 01 2024. [Online]. Available: <http://dx.doi.org/10.5937/fme2402186T>
- [16] P. Krehl and S. Engemann, “August toeppler – the first who visualized shock waves,” *Shock Waves*, vol. 5, no. 1, pp. 1–18, Jun 1995. [Online]. Available: <https://doi.org/10.1007/BF02425031>
- [17] C. Génin, R. Stark, and S. Karl, “Shock system deformation in high mach number rocket nozzles,” in *31st International Symposium on Shock Waves 2*, A. Sasoh, T. Aoki, and M. Katayama, Eds. Cham: Springer International Publishing, 2019, pp. 543–549. [Online]. Available: https://doi.org/10.1007/978-3-319-91017-8_69
- [18] S. B. Verma and C. Manisankar, “Origin of flow asymmetry in planar nozzles with separation,” *Shock Waves*, vol. 24, no. 2, pp. 191–209, Mar 2014. [Online]. Available: <https://doi.org/10.1007/s00193-013-0492-1>
- [19] J. O’stlund and B. Muhammad-Klingmann, “Supersonic flow separation with application to rocket engine nozzles,” *Applied Mechanics Reviews*, vol. 58, no. 3, pp. 143–177, May 2005. [Online]. Available: <https://doi.org/10.1115/1.1894402>
- [20] V. Zmijanović, A. Chpoun, and B. Rasuo, “Flow separation modes and side phenomena in an overexpanded nozzle,” *FME Transactions*, vol. 40, pp. 111–118, 06 2012. [Online]. Available: <https://upsalesiana.ec/ing33ar10r21>
- [21] A. Hadjadj, O. Ben-Nasr, M. Shadloo, and A. Chaudhuri, “Effect of wall temperature in supersonic turbulent boundary layers: A numerical study,” *International Journal of Heat and Mass Transfer*, vol. 81, pp. 426–438, 2015. [Online]. Available: <https://doi.org/10.1016/j.ijheatmasstransfer.2014.10.025>
- [22] R. Zangeneh, “Wall temperature effects on shock unsteadiness in a reattaching compressible turbulent shear layer,” *International Journal of Heat and Fluid Flow*, vol. 92, p. 108876, 2021. [Online]. Available: <https://doi.org/10.1016/j.ijheatfluidflow.2021.108876>
- [23] S. L. Tolentino, J. Mírez, and O. González, “Numerical analysis of the flow field in a planar nozzle for different divergent angles,” *Journal of Mechanical Engineering and Sciences (JMES)*, vol. 16, no. 4, pp. 9241–9252, 2022. [Online]. Available: <https://doi.org/10.15282/jmes.16.4.2022.07.0731>
- [24] R. Arora and A. Vaidyanathan, “Experimental investigation of flow through planar double divergent nozzles,” *Acta Astronautica*, vol. 112, pp. 200–216, 2015. [Online]. Available: <https://doi.org/10.1016/j.actaastro.2015.03.020>
- [25] S. L. Tolentino, J. Mírez, and S. A. Caraballo, “Numerička analiza evolucije udarnog voza u planarnim mlaznicama sa dužinom grla,” *FME Transactions*, vol. 51, no. 4, pp. 595–605, 2023. [Online]. Available: <https://doi.org/10.5937/fme2304595T>
- [26] M. L. Mason, L. E. Putnam, and R. J. Re, *The effect of throat contouring on two-dimensional converging-diverging nozzles at static conditions*. NASA Technical Paper 1704, 1980. [Online]. Available: <https://upsalesiana.ec/ing33ar10r27>
- [27] S. L. Tolentino and J. Mírez, “Efekat dužine grla na obrasce protoka u konusnim mlaznicama van dizajna,” *FME Transactions*, vol. 50, no. 2, pp. 271–280, 2022. [Online]. Available: <https://doi.org/10.5937/fme2201271T>
- [28] B. Wagner and S. Schleichtrien, *Numerical and Experimental Study of the Flow in a Planar Expansion-Deflection Nozzle*. 47th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, 2011. [Online]. Available: <https://doi.org/10.2514/6.2011-5942>
- [29] A. Bourgoing and P. Reijasse, “Experimental analysis of unsteady separated flows in a supersonic planar nozzle,” *Shock Waves*, vol. 14, no. 4, pp. 251–258, Nov 2005. [Online]. Available: <https://doi.org/10.1007/s00193-005-0269-2>
- [30] M. Faheem, A. Khan, R. Kumar, S. Afghan Khan, W. Asrar, and A. M. Sapardi, “Experimental study on the mean flow characteristics of a supersonic multiple jet configuration,” *Aerospace Science and Technology*, vol. 108, p. 106377, 2021. [Online]. Available: <https://doi.org/10.1016/j.ast.2020.106377>
- [31] (ANSYS-Fluent). (2023) Ansys fluent theory guide 2020r1. Ansys Innovation Space. [Online]. Available: <https://upsalesiana.ec/ing33ar10r32>
- [32] F. M. White, *Fluid Mechanics*. McGraw Hill, 2011. [Online]. Available: <https://upsalesiana.ec/ing33ar10r33>
- [33] Y. Egorov, F. R. Menter, R. Lechner, and D. Cokljat, “The scale-adaptive simulation method for unsteady turbulent flow predictions. part 2: Application to complex flows,” *Flow, Turbulence and Combustion*, vol. 85, no. 1, pp. 139–165, Jul 2010. [Online]. Available: <https://doi.org/10.1007/s10494-010-9265-4>

- [34] P. R. Spalart, S. Deck, M. L. Shur, K. D. Squires, M. K. Strelets, and A. Travin, "A new version of detached-eddy simulation, resistant to ambiguous grid densities," *Theoretical and Computational Fluid Dynamics*, vol. 20, no. 3, pp. 181–195, Jul 2006. [Online]. Available: <https://doi.org/10.1007/s00162-006-0015-0>
- [35] F. Menter, M. Kuntz, and R. B. Langtry, "Ten years of industrial experience with the sst turbulence model," *Heat and Mass Transfer*, vol. 4, 01 2003. [Online]. Available: <https://upsalesiana.ec/ing33ar10r36>
- [36] S. Tsan-Hsing, W. L. William, S. Aamir, Y. Zhigang, and Z. Jiang, "A new k- ϵ eddy viscosity model for high reynolds number turbulent flows," *Computers & Fluids*, vol. 24, no. 3, pp. 227–238, 1995. [Online]. Available: [https://doi.org/10.1016/0045-7930\(94\)00032-T](https://doi.org/10.1016/0045-7930(94)00032-T)

GUIDELINES FOR PUBLICATION IN INGENIUS JOURNAL

1. General Information

INGENIUS is a scientific publication of the *Universidad Politécnica Salesiana* of Ecuador, published since January 2007, with a fixed biannual periodicity, specialized in Mechanical Engineering, Electrical Engineering, Electronics, Computer Science and its integration in what is now known as Mechatronics; these lines of action strengthen areas such as automation, control, robotics, among others..

It is a scientific journal, which uses the peer-review system, under double-blind review methodology, according to the publication standards of the Institute of Electrical and Electronics Engineers (IEEE). Compliance with this system allows authors to guarantee an objective, impartial and transparent review process, which facilitates the publication of their inclusion in reference databases, repositories and international indexing.

INGENIUS is indexed in the directory and selective catalog of the Regional Online Information System for Scientific Journals of Latin America, the Caribbean, Spain and Portugal (Latindex), in the Directory of Journals of Open Access DOAJ, In the Information Matrix for the Analysis of Journals, MIAR, In the Ibero-American Network of Innovation and Scientific Knowledge, REDIB and in repositories, libraries and specialized catalogs of Latin America.

The journal is published in a double version: printed (ISSN: 1390-650X) and digital (e-ISSN: 1390-860X), in Spanish, each work being identified with a DOI (Digital Object Identifier System). The articles sent to INGENIUS magazine must comply with the following criteria:

2. Scope and policy

2.1. Theme

Original contributions in Mechanical Engineering, Electrical and Electronic Engineering, Computer Science and its integration in what is now known as Mechatronics, as well as related areas: Automation, Control, Domotics, Robotics in their different fields of action and all those related disciplines with the same central theme.

All the work carried out by national or foreign researchers may be published once they meet the required scientific quality criteria.

2.2. Contributions

INGENIUS Journal preferably publishes articles related to empirical research, and also reports of technological development, proposals for models and innovations, products for the elaboration of graduate and postgraduate thesis that contribute to the field of science and technology, as well as select revisions of literature. (state-of-the-art).

- **Research:** 5,000 to 6,500 words of text, including title, abstracts, descriptors, charts and references.
- **Reports:** 5,000 to 6,500 words of text, including title, abstracts, charts and references.
- **Reviews:** 6,000 to 7,000 words of text, including charts and references. Current, selective and justified references, would be specially valued from among 40 works

The INGENIUS Journal publishes original and unpublished works written in Spanish and English, they may not have been published

through any printed or electronic media, nor be in the process of arbitration or publication.

Every article will be subjected to a rigorous arbitration process; the evaluation of the article will be made according to criteria of originality, relevance, relevance, contributions, scientific rigor and compliance with established editorial guidelines.

Being an arbitrated publication, the Editorial Board approves its publication based on the concept of specialized pairs. The reception of a document does not imply commitment of publication.

It is essential to present a letter of presentation and grant of rights which can be downloaded from: [urlhttps://goo.gl/ZNkMRD](https://goo.gl/ZNkMRD).

Contributions must be exclusively sent and through the OJS (Open Journal System) [urlhttps://goo.gl/JF7dWT](https://goo.gl/JF7dWT). In which all authors must previously register as a user. For any consultation of the procedure you should contact:

revistaingenius@ups.edu.ec,
jcalles@ups.edu.ec ó
mquinde@ups.edu.ec.

To promote diversity in publications, the author(s) may not publish more than one (1) article per issue, nor in consecutive issues. In order for the author(s) to submit their research again to Ingenius, a minimum of 3 published issues must have elapsed.

3. Presentation and structure of the manuscripts

For those works that are empirical investigations, the manuscripts will follow the IMRDC structure (Introduction, Materials and Methods, Results and Discussion and Conclusions), being optional the Notes and Supports. Those papers that, on the contrary, deal with reports, studies, proposals and reviews may be

more flexible in their epigraphs, particularly in material and methods, analysis, results, discussion and conclusions. In all typologies of works, references are mandatory.

Articles may be written on Microsoft Word (.doc or .docx) or L^AT_EX(.tex). The template to be used can be downloaded from the journal's website, a, [urlhttps://goo.gl/gtCg6m](https://goo.gl/gtCg6m), while for L^AT_EX in [urlhttps://goo.gl/hrHzzQ](https://goo.gl/hrHzzQ), it is necessary that the file be anonymised in Properties of File, so that the author(s) ID is not displayed.

Figures, Graphs and/or Illustrations, as well as Charts shall be numbered sequentially including an explanatory description for each. The equations included in the article must also be numbered; the figures, charts and equations must be cited in the text.

Use space after point, commas and question marks.

Use “enter” at the end of each paragraph and title heading. Do not use .^{enter}.^{anywhere} else, let the word processor program automatically break the lines.

Do not center headings or subheadings as they should be aligned to the left.

Charts must be created in the same program used for the document body, but must be stored in a separate file. Use tabs, not spaces, to create columns. Remember that the final size of printed pages will be 21 x 28 cm, so the tables must be designed to fit the final print space.

3.1. Structure of the manuscripts

3.1.1. Presentation and cover letter

1. **Título (español) / Title (inglés):** Concise but informative, in Spanish on the front line and in English on the second, when the article is written in Spanish and vice versa if it is written in English.

2. **Authors and affiliations:** Full name and surname of each author, organized by order of priority and their institutional affiliation with reference to the end of the first sheet, where it must include: Dependency to which belongs within the institution, Institution to which he/she belongs, country, ORCID. A maximum of 5 authors will be accepted, although there may be exceptions justified by the complexity and extent of the topic.
 3. **Abstract (Spanish) / Abstract (English):** It will have a maximum extension of 230 words, first in Spanish and then in English. : 1) Justification of the topic; 2) Objectives; 3) Methodology and sample; 4) Main results; 5) Main conclusions.
 4. **Keywords (Spanish) / Keywords (English):** 6 descriptors must be presented for each language version directly related to the subject of the work. The use of the key words set out in UNESCO's Thesaurus will be positively valued.
 5. **Presentation (Cover Letter):** A statement that the manuscript is an original contribution, not submission or evaluation process in another journal, with the confirmation of the signatory authors, acceptance (if applicable) of formal changes in the manuscript according to the guidelines and partial assignment of rights to the publisher, according to the format established in: <<https://goo.gl/ZNkMRD>>
- nal affiliation with reference to the end of the first sheet, where it must include: Dependency to which belongs within the institution, Institution to which he/she belongs, country, ORCID. A maximum of 5 authors will be accepted, although there may be exceptions justified by the complexity and extent of the topic.
3. **Abstract (Spanish) / Abstract (English):** It will have a maximum extension of 230 words, first in Spanish and then in English. : 1) Justification of the topic; 2) Objectives; 3) Methodology and sample; 4) Main results; 5) Main conclusions.
 4. **Keywords (Spanish) / Keywords (English):** 6 descriptors must be presented for each language version directly related to the subject of the work. The use of the key words set out in UNESCO's Thesaurus will be positively valued.
 5. **Introduction:** It should include the problem statement, context of the problem, justification, rationale and purpose of the study, using bibliographical citations, as well as the most significant and current literature on the topic at national and international level.
 6. **Material and methods:** It must be written so that the reader can easily understand the development of the research. If applicable, it will describe the methodology, the sample and the form of sampling, as well as the type of statistical analysis used. If it is an original methodology, it is necessary to explain the reasons that led to its use and to describe its possible limitations.
 7. **Analysis and results:** It will try to highlight the most important observations, describing, without making value judgments, the material and methods used. They will appear in a logical sequence

3.1.2. Manuscript

1. **Título (español) / Title (inglés):** Concise but informative, in Spanish on the front line and in English on the second, when the article is written in Spanish and vice versa if it is written in English.
2. **Authors and affiliations:** Full name and surname of each author, organized by order of priority and their institutional affiliation with reference to the end of the first sheet, where it must include: Dependency to which belongs within the institution, Institution to which he/she belongs, country, ORCID. A maximum of 5 authors will be accepted, although there may be exceptions justified by the complexity and extent of the topic.

in the text and the essential charts and figures avoiding the duplication of data.

8. **Discussion and Conclusions:** It will summarize the most important findings, relating the observations themselves to relevant studies, indicating contributions and limitations, without adding data already mentioned in other sections. It should also include deductions and lines for future research.
9. **Supports and acknowledgments (optional):** The Council Science Editors recommends the author (s) to specify the source of funding for the research. Priority will be given to projects supported by national and international competitive projects.
10. **The notes (optional):** will go, only if necessary, at the end of the article (before the references). They must be manually annotated, since the system of footnotes or the end of Word is not recognized by the layout systems. The numbers of notes are placed in superscript, both in the text and in the final note. The numbers of notes are placed in superscript, both in the text and in the final note. No notes are allowed that collect simple bibliographic citations (without comments), as these should go in the references.
11. **References:** Bibliographical citations should be reviewed in the form of references to the text. Under no circumstances should references mentioned in the text not be included. Their number should be sufficient to contextualize the theoretical framework with current and important criteria. They will be presented sequentially in order of appearance, as appropriate following the format of the IEEE.

3.2. Guidelines for Bibliographical references

Journal articles:

- [1] J. Riess, J. J. Abbas, "Adaptive control of cyclic movements as muscles fatigue using functional neuromuscular stimulation". IEEE Trans. Neural Syst. Rehabil. Eng vol. 9, pp.326–330, 2001. [Online]. Available: <https://doi.org/10.1109/7333.948462>

Books:

- [1] G. O. Young, "Synthetic structure of industrial plastics" in *Plastics*, 2nd ed., vol. 3, J. Peters, Ed. New York: McGraw–Hill, 1964, pp. 15–64.

Technical reports:

- [1] M. A. Brusberg and E. N. Clark, "Installation, operation, and data evaluation of an oblique–incidence ionosphere sounder system," in "Radio Propagation Characteristics of the Washington–Honolulu Path," Stanford Res. Inst., Stanford, CA, Contract NOBSR–87615, Final Rep., Feb. 1995, vol. 1

Articles presented in conferences (unpublished):

- [1] Vázquez, Rolando, Presentación curso "Realidad Virtual". National Instruments. Colombia, 2009.

Articles of memories of Conferences (Published):

- [1] L. I. Ruiz, A. García, J. García, G. Taboada. "Criterios para la optimización de sistemas eléctricos en refinerías de la industria petrolera: influencia y análisis en el equipo eléctrico," IEEE CONCAPAN XXVIII, Guatemala 2008.

Thesis:

- [1] L.M. Moreno, "Computación paralela y entornos heterogéneos," Tesis doctoral, Dep. Estadística, Investigación Operativa y Computación, Universidad de La Laguna, La Laguna, 2005.

Guidelines:

- [1] IEEE Guide for Application of Power Apparatus Bushings, IEEE Standard C57.19.100–1995, Aug. 1995.

Patents:

- [1] J. P. Wilkinson, “Nonlinear resonant circuit devices,” U.S. Patent 3 624 125, July 16, 1990.

Manuals:

- [1] Motorola Semiconductor Data Manual, Motorola Semiconductor Products Inc., Phoenix, AZ, 1989.

Internet resources:

- [1] E. H. Miller, “A note on reflector arrays” [Online]. Available. <https://goo.gl/4cJkCF>

3.3. Epigraphs, Figures and Charts

The epigraphs of the body of the article will be numbered in Arabic. They should go without a full box of capital letters, neither underlined nor bold. The numbering must be a maximum of three levels: 1. / 1.1. / 1.1.1. At the end of each numbered epigraph will be given an enter to continue with the corresponding paragraph.

The charts must be included in the text according to order of appearance, numbered in Arabic and subtitled with the description of the content, the subtitle should go at the top of the table justified to the left.

Figures can be linear drawings, maps or black and white halftone or color photographs in 300 dpi resolution. Do not combine photographs and line drawings in the same figure.

Design the figures so that they fit eventually to the final size of the journal 21 x 28 cm. Make sure inscriptions or details, as well as lines, are of appropriate size and thickness so that they are not illegible when they are reduced to their final size (numbers, letters and symbols must be reduced to at least 2.5 mm in height After the illustrations have

been reduced to fit the printed page). Ideally, the linear illustrations should be prepared at about a quarter of their final publication size.

Different elements in the same figure should be spelled a, b, c, etc.

Photographs should be recorded with high contrast and high resolution. Remember that photographs frequently lose contrast in the printing process. Line drawings and maps should be prepared in black.

The text of the figures and maps must be written in easily legible letters.

If the figures have been previously used, it is the responsibility of the author to obtain the corresponding permission to avoid subsequent problems related to copyright.

Each figure must be submitted in a separate file, either as bitmap (.jpg, .bmp, .gif, or .png) or as vector graphics (.ps, .eps, .pdf).

4. Submission process

The manuscript must be sent through the OJS system of the journal, <<https://goo.gl/JF7dWT>>, the manuscript should be uploaded as an original file in .pdf without author data and anonymized according to the above; In complementary files the complete manuscript must be loaded in .doc or .docx (Word file), that is to say with the data of the author (s) and its institutional ascription; Also the numbered figures should be uploaded in independent files according to the corresponding in the manuscript (as bitmap .jpg, .bmp, .gif, or .png or as vector graphics .ps, .eps, .pdf). It is also obligatory to upload the cover letter and grant of rights as an additional file.

All authors must enter the required information on the OJS platform and only one of the authors will be responsible for correspondence.

Once the contribution has been sent the system will automatically send the author for correspondence a confirmation email of receipt

of the contribution.

5. Editorial process

Once the manuscript has been received in OJS, a first check by the editorial team of the following points:

- The topic is in accordance with the criteria of the journal.
- Must have the IMRDC structure.
- Must be in the INGENIUS format.
- Must use the IEEE citation format.
- All references should be cited in the text of the manuscript as well as charts, figures and equations.
- The manuscript is original; for this, software is used to determine plagiarism.

The assessment described above can take up to 4 weeks.

If any of the above is not complete or there is inconsistency, an email will be sent to the author to make the requested corrections.

The author will make the corrections and resend the contribution through an email in response to the notification and will also upload the corrected manuscript into OJS supplementary files.

The editorial team will verify that the requested corrections have been incorporated, if it complies, the manuscript will start the second part of the process that may be followed by the author through OJS, otherwise the author will be notified and the manuscript will be archived.

The second phase of the process consists of the evaluation under the methodology of double-blind review, which includes national and foreign experts considering the following steps:

- The editor assigns two or more reviewers for the article.
- After reviewing the article, the reviewers will submit the evaluation report with one

of the following results.

- Publishable
- Publishable with suggested changes
- Publishable with mandatory changes
- Non publishable
- The editor once received the evaluation by the reviewers will analyze the results and determine if the article is accepted or denied.
- If the article is accepted, the author will be notified to make corrections if required and the corresponding editorial process will be continued.
- If the article is denied, the author will be notified and the manuscript will be archived.
- In the two previous cases the result of the evaluation of the reviewers and their respective recommendations will be sent.

The second phase of the process lasts at least 4 weeks, after which they will be notified to the author giving instructions to continue with the process.

6. Publication

The INGENIUS Journal publishes two issues per year, on January 1st and July 1st, so it is important to consider the dates for sending the articles and their corresponding publication. Articles received until October will be considered for the January publication and those received until April for the July publication.

7. Information on the Use of Artificial Intelligence

Should artificial intelligence be used at any stage of the research presented in the article, authors are required to clearly highlight this in the cover letter associated with the article, specifying the section or sections where artificial intelligence has been used. The purpose

of this requirement is to inform readers about the sections where this technology has been employed, providing greater transparency and understanding of its application in the presented research.

INGENIUS, Revista de Ciencia y Tecnología, recognizes the importance of maintaining

high ethical standards in scientific research, particularly in the use of artificial intelligence (AI).

The decision to accept a publication that has utilized artificial intelligence rests at the discretion of the editorial team.

UNIVERSIDAD POLITÉCNICA SALESIANA DEL ECUADOR

Juan Cárdenas Tapia, sdb,
Rector

©Universidad Politécnica Salesiana
Turuhuayco 3-69 y Calle Vieja
Postal code 2074
Cuenca, Ecuador
Teléfono: (+593 7) 205 00 00
Fax: (+593 7) 408 89 58
Email: srector@ups.edu.ec

Exchange

Exchange with other periodicals is accepted.

Address:

Secretaría Técnica de Comunicación
Universidad Politécnica Salesiana
Turuhuayco 3-69 y Calle Vieja
Postal code 2074
Cuenca, Ecuador
Phone: (+593 7) 205 00 00 Ext. 1182
Fax: (+593 7) 408 89 58
Email: rpublicas@ups.edu.ec
www.ups.edu.ec
Cuenca – Ecuador

INGENIUS, Journal Science of Technology,
Issue 33
january/june 2025
John Calle Sigüencia, Editor in chief
revistaingenius@ups.edu.ec

Printed

Centro Gráfico Salesiano: Antonio Vega Muñoz 10-68 y General Torres.
Phone: (+593 7) 283 17 45
Cuenca – Ecuador
Email: centrograficosalesiano@lms.com.ec

OTHER PERIODIC PUBLICATIONS OF THE UNIVERSITY

UNIVERSITAS, Journal of Social and Human Sciences.

LA GRANJA, Journal of the Sciences.

ALTERIDAD, Journal of Education.

RETOS, Journal of Administration Sciences and Economics.

UTOPIA, University Youth Ministry Magazine.

SOPHIA, Collection of Philosophy of Education.



**ABYA
YALA** | UNIVERSIDAD
POLITÉCNICA
SALESIANA

